

Other people's nuclear weapons

Week-end diplomacy seems to have won Britain a prolongation of its role as a nuclear power, but the time has come to ask when and how this will be attenuated.

THE British government's case for maintaining nuclear forces is far from simple and far from clear. Historically, the British stumbled into making nuclear weapons because, in the late 1940s, it seemed to the then Labour government that this was the natural course to follow in a substantial and technologically advanced country. Over the years, circumstances have changed. The range of Britain's international interests has shrunk to an extent that could not have been foreseen forty years ago, while the country's industrial economy is no longer large enough to sustain the independent development and construction of strategic delivery systems.

The obvious parallel with the present time is that in the early 1960s, when the then Prime Minister, Mr Harold Macmillan (now Lord Stockton), persuaded President John F. Kennedy that, if the air-launched nuclear bomb called Skybolt was not available (because it did not function as designed), the British government should be allowed to purchase Polaris submarines instead. But why? In one sense, it is merely right and proper that a government which is a member of an alliance such as the North Atlantic Treaty Organisation (NATO) should make an equitable contribution to its strategic forces; indeed, the British Polaris submarines are operated within NATO guidelines, but with the understanding that they might be withdrawn, and operated independently, if some vital national interest were involved.

But what could those circumstances be? Is not the United States fully committed to the defence of Europe not merely through the formal treaty on which NATO rests but also by the presence both of troops and of strategic missiles on European soil? The simple answer is that these assurances have been insufficient for a succession of British governments, all of which have recognized that circumstances could arise when the United States might shrink from using its strategic nuclear forces, risking retaliation in the process, for the sake of some ill-understood European interest. These dark suspicions have been strengthened by the news from Reykjavik last month.

Negotiations

Whether Britain and France will be able indefinitely to continue in that frame of mind is another matter. Everything will depend on what happens in the negotiations on strategic arms still continuing at Geneva. After Reykjavik, and the recognition by the two most powerful people in the world that even they cannot magic out of thin air lasting agreements on difficult questions in a dozen hours of conversation, it is inevitable that strategic arms control will be regarded as a matter for the long haul; the bilateral negotiations at Geneva will (with luck) be prolonged into the next decade, and more attention will be paid to other aspects of the relationship between the major powers. But the time will come when, again with luck, there will be another proposal that the balance between both strategic missiles and missiles of intermediate range should be struck at a lower level than at present. At some stage, the point will (with luck, again) arrive at which the British and French nuclear forces stand out conspicuously. Then the superpowers will be pressing to see these independent nuclear forces cut back to size. The United

States, having helped Britain to build Trident submarines, will be wanting to see them scuttled. That will be an awkward conversation for some future weekend at Camp David.

The British position on this foreseen dilemma is, fortunately, explicit. In the first place, the submarine missiles are regarded as strategic weapons, contributing now to the strategic deterrent of NATO as a whole but also intended as a means by which Britain might threaten even a more powerful potential adversary with an unacceptable degree of retaliatory damage. Theoretically, the French regard for nuclear weapons is very much the same; they might not be sufficient to influence events on the global stage, but they would surely 'tear an arm' from an attacker, and give him pause. But the British government has also gone so far as to consider when it would be proper to think of limiting its nuclear forces, saying that this would be acceptable as part of a more general agreement for the substantial reduction of nuclear forces everywhere.

Unfortunately, this undertaking is not nearly as clear as it might be. The British government argues that, because its weapons are strategic, their number should not be affected by an agreement on the numbers of superpower nuclear weapons deployed in or against Europe, which makes sense. But the prospect that there may be very substantial reductions of European nuclear forces can only strengthen the British government's unspoken fears that Europe may one day be left to its own devices, whence in part Mrs Thatcher's visit to Washington at the weekend. The snag is that the British position is vague on what would happen if there were also a substantial agreement on strategic forces might be held to provide a licence for conventional wars in Europe, in which West European states would be disadvantaged. So British governments will be tempted to hang on to their nuclear weapons through thick and thin.

The lessons to be learned from these tortuous arguments are several. First, Britain's susceptibilities about its nuclear weapons mirror those of other European states about other aspects of their arrangements for defence. British nuclear weapons would indeed be an impediment to sudden and drastic reductions of nuclear arms along the lines fancied at Reykjavik, but so would be the political protestations of other European states, from which it follows that the superpowers must aim in their negotiations at a balanced package of arms control agreements covering not merely strategic and intermediate nuclear missiles but conventional forces and even the temper of East-West relations. On that score, the Europeans have a case. But the United States also has a right to expect that the British and French nuclear weapons will not be impediments to reasonable arms agreements that may in future be reached with the Soviet Union. Indeed, it would have been natural that President Reagan should have broached the question with Mrs Thatcher at the weekend, while confirming that Trident should succeed Polaris. And, given governments' temptations to hang on to weapons in the face of arguments to the contrary, there is the strongest case for asking that the British government should now be prepared publicly to define the circumstances in which it will let go. France, which is a harder case, should be persuaded also to follow suit. □



Carnegie's misdiagnosis

A new study of undergraduate education misses an opportunity to tell what is really wrong.

US ACADEMICS and their followers, in one of their periodic fits of introspection, have been giving pious attention to a report from the Carnegie Foundation for the Advancement of Teaching on the state of higher education and, in particular, to its judgement that, "driven by careerism and overshadowed by graduate and professional education, many of the nation's colleges and universities are more successful in credentialing (*sic*) than in providing a quality education for their students". There will be few who disagree with that, but the foundation's study, *College — the undergraduate experience in America*, by Mr Ernest Boyer, is most notable for what it has left unsaid.

Carnegie is, naturally, right to be troubled by the lack of individual attention given to undergraduates and to sympathize with them in the hard choice of a course of study. Then few will quarrel with Boyer's plea for better personal advice to students, especially during the difficult transition from high-school to college. And nobody will argue against the case for higher standards of literacy. But the argument hangs too much on the subjective experiences of undergraduates and teachers, as reported and, often touchingly, recorded. This preoccupation with tensions between "individual interests and shared concerns" means that some of the urgent questions about the future of higher education in the United States are mis-stated.

One is the pattern of the curriculum. US universities have always required less specialization than have universities elsewhere, and deserve credit for it. Their accessibility is unmatched anywhere in the world. But the worry now is that things have gone too far, and that US undergraduates have too much latitude to choose from a curriculum increasingly resembling a supermarket for credit courses, with students free to experiment as they please until the required number of credits is reached. The Carnegie report rightly raises this important issue, with which William J. Bennett, the US Secretary of Education, startled some of those at Harvard University's 350th birthday celebrations earlier this year, taking Harvard as his inappropriate example. Carnegie's remedy for the laundry-list approach to education is to call for a core curriculum. The need for some high-level coordination cannot be denied. But if Carnegie wishes to be taken seriously, it must devise a more persuasive core than the amalgam now put forward; an "integrated core" that would include language, art, history, institutions, nature ("the ecology of the planet"), work ("the value of vocation") and identity ("the search for meaning"). All this, but no general science course?

The Carnegie study also draws attention to the conflicts that can arise in research universities between the demands of teaching and of research. The story is familiar: because academic prowess is measured largely by published research papers, faculty members devote their best efforts to excellence in research, sometimes at the expense of teaching. Boyer's suggestion that research universities should establish a new rank of "distinguished teaching professor", extending status and salary incentives to outstanding teachers, might serve well in some colleges and universities, but ignores the frequency with which excellence in teaching and in research go together, hand in glove. In the end, the research universities will have to deal with this internal conflict internally, by requiring that faculty members are not so distracted by external interests and commitments as to impede their discharge of their academic responsibilities, to students as well as to research. This goal will not be easily won, but the research universities eager that their reputations should survive will give it more attention than has been the habit in recent years. The professorial status of good teachers is a palliative only.

The Carnegie Foundation would also have done education a

greater service if it had tackled some of the problems of higher education that are beyond the control of the universities themselves. The most immediate threat is the accumulated effect of the dereliction of university research facilities and cutbacks in support for students. Many of the institutions that have served the United States well in the past, the less well-known universities that dominate the Boyer's survey in particular, are now ominously stretched. Government officials are fond of pointing out that support for basic research at universities has increased in recent years, which is true. But the increase followed a long decline during the 1970s that has not yet been made good. Decaying facilities and equipment are depressingly common on all but the most successful campuses. Federal support for university research facilities dropped from \$404 million in 1966 to a low point of just \$37 million in 1981 (1981 prices). Support for graduate students in the sciences has fallen by about a half since 1969. The efforts of the National Science Foundation to promote excellence in science education, all but abolished by President Reagan at the start of his first term of office, are improving, but slowly. Small wonder that the proportion of bachelor's degrees in physical and biological sciences has fallen since the late 1970s, with even engineering starting to tail off in 1985. Yet fewer undergraduates now means fewer faculty members in ten years' time, less research and, probably, less teaching. Unless the federal government decides to act, these are the circumstances that will condition some future Carnegie report on the parlous condition of higher education in the United States at the end of the decade. □

Fire without smoke

Two retracted papers do not make a scandal, but too much reticence about the circumstances may.

DANA-Farber Cancer Institute has a problem (see page 197). It has announced that two papers describing a T-cell activating substance called interleukin-4A are being retracted from the journals in which they were published. This is certainly not the first time that scientific papers have been withdrawn, and it will not be the last. It is very much to the credit of Dana-Farber that the retractions were issued as soon as the errors in the data were discovered.

Where Dana-Farber's problem lies is in the way it described — or more properly failed to describe — the reasons for the retraction. Obviously there are many reasons why papers may have to be retracted. Cell culture contamination, reagent impurities, inadvertent miscoding of results can all be discovered, long after the fact, even in the most careful and thorough laboratories. But Dana-Farber has not invoked any of these readily believable explanations, choosing to make the blunt statement that an investigation is under way, and that retraction letters have gone out. This approach can only leave a cloud of suspicion hanging over the institute. What sort of investigation is it that can immediately conclude that the data upon which a paper is based are incorrect and not reproducible without saying how that fact is known?

Unhappily, there have been all too many recent instances in which falsification of data has brought laboratories into trouble. By declining to comment about its investigation, Dana-Farber, perhaps unwittingly, merely encourages speculation. Officials at Dana-Farber will correctly say that it is unfair to make allegations that could have a profound influence on a scientist's career before they are completely investigated. But silence is not the answer. Surely it is possible to say what first aroused suspicions about the data without jeopardizing future investigation? Surely retractions with no explanation will focus interest on the case longer than a simple statement about the circumstances as they are at present known? The scientific community should be credited with the good sense to understand the distinction between an investigation and a final judgement. □

Research publishing

Interleukin disappears from the literature

Boston

Two papers from Dr Ellis Reinherz's laboratory at Dana-Farber Cancer Institute that announced the discovery and properties of a T-cell activating substance termed interleukin-4A (IL-4A) have been retracted by the authors. The reason given for the retraction is that the data are incorrect and are not reproducible. The authors, in other words, no longer claim to have proof of the existence of IL-4A originally hailed as the chemical messenger that amplifies the body's immune response by turning on resting T cell's, through its interaction with the T11 protein at the surface of the T cell.

The papers being retracted were published in *Science* (231, 1118; 1986 by Claudio Milanese, Neil E. Richardson and Ellis L. Reinherz) and the *Journal of Experimental Medicine* (163, 1583; 1986 by Claudio Milanese, Robert E. Siliciano, Reinhold E. Schmidt, Jerome Ritz, Neil F. Richardson and Ellis L. Reinherz). Letters of retraction are being sent to both journals.

An *ad hoc* committee of scientists from both within and outside Dana-Farber is investigating why the papers need to be retracted. Both Reinherz and David Kiszka, director of research at the institute, refused to comment on the reasons the data became suspect, how the data were proved incorrect, or why a retraction of the articles was necessary even before an investigation was completed. It is possible that litigation may result, according to Reinherz; but a final decision will be up to Dana-Farber and its parent Harvard Medical School or perhaps the National Institutes of Health (NIH), Kiszka says.

Kiszka refused to comment on the process of data review at Dana-Farber or in the Reinherz laboratory. Dr Sheldon Wolff of the New England Medical Center, says any person in a laboratory understands there will be efforts to reproduce all published data. There is no way the principal scientist can be expected to stand at the elbow of everyone working in a laboratory, says Wolff. The work of Reinherz is unanimously described as "impeccable" by others in the field.

The withdrawal of the IL-4A papers means there is now only one well-studied candidate for a T11 ligand, a molecule termed LFA3. The data supporting the existence and activity of this molecule are so convincing, says Dr Peter Beverley at the Imperial Cancer Research Fund in London, that "a number of people were sceptical of the dramatic announcement"

of IL-4A. According to Dr Stephen Shaw of NIH, a principal investigator of LFA3, there is much interest in finding the one or more molecules that interact with T11, also called CD2. T11 is a marker on T cells thought to be involved in adhesion of lymphocytes and triggering of T-cell activation. IL-4A was described as soluble, of

low molecular weight, and responsible for triggering resting T cells. LFA3 is a much larger cell-bound molecule responsible for lymphocyte adhesion.

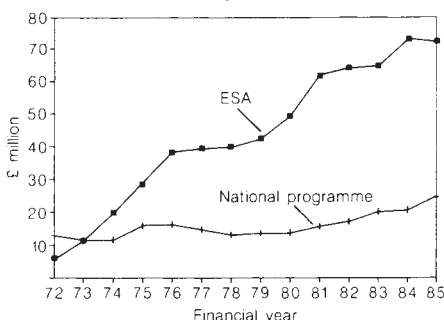
Shaw says he originally accepted the IL-4A data because they came from the Reinherz laboratory and because there is room on the T11 molecule for two different recognition sites. He also thought IL-4A could be a soluble fragment of LFA3. Even though there is no longer proof of the molecule termed IL-4A, he says, other ligands for the "popular and trendy" T11 molecule may exist, including (he speculates) a soluble form of LFA3.

Elizabeth Collins

British space

Waiting for an increased budget

THE British National Space Centre (BNSC) this week celebrates its first anniversary still awaiting the government's response to its 15-year plan for Britain's presence in space. The blueprint, if accepted, will mean a substantial increase in the £108 million a year that Britain now



BNSC seeks more funds for space research and a narrowing of the gap between international and domestic projects.

spends on international and domestic space projects.

Britain is under pressure from the United States and Europe to invest heavily. It was the need to respond sensibly to an invitation to participate in the US Space Platform that inspired the creation of BNSC (see *Nature* 318, 305; 1986).

A ministerial council meeting of the European Space Agency (ESA) in Rome in 1985 generated pressure from Europe. The meeting approved in principle a number of new programmes that would require members' contributions to be increased by 70 per cent in the next three years. France, for example, which already spends five times as much as Britain, has decided to increase its budget by 40 per cent, bringing the total annual expenditure on space to £700 million.

Roy Gibson, director general of BNSC, says "the European scene is going on, whether we are ready or not", and that BNSC must have a positive response from government within the next few weeks if proper preparations are to be made for the

ESA ministerial council meeting next June at which the final touches to the ESA programme, some lasting more than a decade, will be made.

"Last year we weren't too worried about this deadline in June 1987. . . now we're getting a big edgy because we need these decisions in order to give us the best bargaining position inside ESA."

The British space plan, he says, will require substantial international investment and the domestic budget must be increased to ensure industrial back-up to exploit the new technologies. Much of the money for Britain's contribution to the space platform project, Gibson believes, must come from the government. Although the private sector may eventually benefit, it would find it hard to justify a commercial investment at this stage. "I think it would be very wrong to tell government that it's going to make a lot of money straight away."

BNSC wants the government to commit itself to plans that will span more than a decade and to allocate money at least three years ahead to allow industry to respond and plan accordingly.

To that end, the BNSC plan details the principal areas in which the United Kingdom must develop. These are space science, remote sensing, microgravity, telecommunications (with particular emphasis on mobile communications on land, air and sea), space station, ground/test/reception facilities, data processing and enabling technologies.

The present budget is culled from the space funds of the Department of Trade and Industry, the Ministry of Defence, the Natural Environment Research Council and the Science and Engineering Research Council. The 280 members of staff are spread between London, the Rutherford Appleton Laboratory in Oxfordshire and Farnborough in Hampshire. The next stage would be to create a permanent role for BNSC and perhaps a permanent home.

Bill Johnstone

Soviet accusation

Plagiarism charges levelled

Two leading members of the Academy of Sciences of the Byelorussian SSR have accused a US professor, Dr Hollis B. Chen, of plagiarism. Writing in the Russian weekly *Literaturnaya Gazeta*, Dr Mikolaj Barysevic, president of the Byelorussian Academy, and Dr Barys Sciapanau, a three-times State Prize winner, have accused Chen of pirating chapters 3–6 of his textbook *Theory of Electromagnetic Waves — a Coordinate-Free Approach* (McGraw Hill, 1985). This material, they claim, was lifted without acknowledgement from two monographs by Dr Fiodar Fiodarau, son of the Byelorussian writer Janka Maur. Fiodarau is a leading light of the Byelorussian Academy, and is secretary of its section of physical, mathematical and technical sciences.

According to Barysevic and Sciapanau, the “coordinate-free” method was pioneered by Fiodarau more than 30 years ago. On the basis of this “fruitful approach”, they say, he has published six monographs and “hundreds of scientific articles”. Chen’s work, they allege, is based on two of three works, the monograph *Optics of Anisotropic Media* (Minsk, 1958) and the book *Theory of Gyrotropy* (Minsk, 1976). Both works are fairly obscure, having been published, the critics admit, “a fairly long time ago, in a small edition, and from an obscure scientific press (*Navuka i Technika*, Minsk).

Their accusations may be summarized under four points.

- Chen, they say, claims on the blurb that his method is new, and that the book “for the first time treats the essence of the subject fully on a coordinate-free basis”;
- Chen’s bibliography makes no reference to Fiodarau in its more than 200 items;
- Chapters 3–6 of Chen’s book are almost completely copies from Fiodarau, and, in particular, the diagrams have been lifted unchanged;
- Chen’s book, “contrary to academic-

Chen’s statement

“This book is intended as a text for seniors and first-year graduate students, not Research monograph. The development of the book relies heavily on my previous research work on waves in plasma and the published work of the others as stated in the preface. I have never claimed the creation of the new method. In fact, the vectors and dyads were invented in 1881 by Professor J.W. Gibbs of Yale University. Unfortunately, I do not read Russian and most of my references were in English. This is because the focus was on developing a strong textbook in electromagnetic waves rather than a comprehensive research monograph”. □

tradition”, does not give references, within the body of the text, to the works mentioned in the footnotes. This, say the critics, is to avoid revealing how much material is derived from Fiodarau.

The complaint ends by listing the many learned societies to which Chen belongs, in terms that leave no doubt that, in the opinion of the critics, he should be expelled as “one of the most cynical and shameless plagiarists in the history of science”.

Examination of the book does not appear to substantiate these allegations. Chen’s work, a textbook for final-year undergraduate and graduate students, does not claim any novelty of method, although his preface does speak of its

being an “alternative approach”, a coordinate-free approach, and notes that “the organization of the book is somewhat different from that normally found in books on electromagnetics”. His diagrams, in the disputed Chapters 3–6 in particular, seem standard vectorial representations of such phenomena as refraction and reflection. But Chen in fact makes little claim to novelty and was not, he said, responsible for the jacket. In a statement to *Nature* last week he says that he had never claimed the creation of a new method, and maintains that he had relied heavily on the previous work of others. The book is not, he emphasizes, a research monograph but a university text. He therefore makes little reference to work published in languages other than English. He has no knowledge, he stresses, of Russian.

Vera Rich

Australian universities

Going private pleasantly?

Sydney

PLANS for a private university that would break the public monopoly in higher education are attracting strong opposition from academics, in particular university staff associations and teachers’ groups. The address of the Bond University of Technology is Miami, Surfers Paradise, on Queensland’s Gold Coast, one of Australia’s most pleasant areas. It will open its doors in 1989 to a first intake of 1,000 out of a planned total of 10,000 students, many of whom will be from overseas, particularly from Japan and South-East Asia. But the organizers hope that at least half of the students will come from Australia’s free government tertiary institutions.

The university is the brainchild of businessman Alan Bond, of America’s Cup fame. Funds for the university will come equally from the Bond Corporation and a Japanese development company represented by Dr Bungo Ishizaki of the Osaka University of Commerce. The university will be run as a non-profit body with operating costs covered entirely by student fees. The backers intend to make their investment pay by capitalizing on products and inventions arising from research carried out in the university.

High salaries will be offered to attract talented staff and the position of vice-chancellor is being advertised at a salary of A\$150,000, substantially higher than that for a similar post at an existing university.

Criticisms of the Bond University include doubts about the standard of its courses and whether or not the range of subjects to be offered warrants the title of university. The main reason for opposition seems to be the threat it poses to Australia’s egalitarian education system by creating an elite of those who can afford to pay. One fear is that, once the institution is established,

it may be successful in pressing the government for assistance, and that with the proliferation of such private institutions, existing establishments will be left with an ever smaller cake to cut up. Others worry that the institution will become a last resort for those not bright enough to gain admittance elsewhere. And there is concern over the destabilizing effect of the high salaries on the rigid salary structure of the government institutions.

Supporters feel that the quality of present universities is suffering because of the restrictions that accompany government funding. Dr Ian Bassett, a senior lecturer in physics at the University of Sydney, believes that salary flexibility is vital to attract outstanding people. Present arrangements of the Commonwealth Tertiary Education Commission (CTEC) prevent universities from offering higher salaries even if the extra comes from private sources. When the University of Sydney recently gave its vice-chancellor a few thousand dollars extra, the same amount was taken out of the following year’s grant. And the present rigidity in salaries fails to recognize that it is more difficult to attract talented staff to some disciplines than to others.

In defence of its own plan, the Bond Corporation points out that it is in line with the government’s policy of encouraging more contact between academics and industry. Furthermore, with a rapidly growing population of 370,000, the Gold Coast region is the seventh largest population centre in Australia and the only one of its size without a university. The present climate of restraint in government spending is unlikely to give the region another chance to get one say the supporters of the Bond plan.

Charles Morgan

Nuclear energy

India joins enrichment club

New Delhi

INDIA has developed the capability to enrich uranium, according to Dr Raja Ramanna, chairman of the Atomic Energy Commission (AEC); he did not reveal the size of the enrichment plant nor the amount of U-235 produced. Ramanna was speaking at a press conference and reacting to US press reports that Pakistan is now able to make bomb-grade uranium.

For the past four years, the Bhabha Atomic Research Centre (BARC) has been operating a gas centrifuge pilot plant at Trombay near Bombay. Its operation and progress have been closely guarded secrets until now. Unlike Pakistan, which relied on secret imports of design and parts for its Kahuta enrichment facility,

BARC is believed to have designed the plant itself.

There are rumours that, based on the pilot facility, BARC is setting up a full-scale gas centrifuge plant near Mysore in Karnataka state in South India. But Iyengar, in an interview on All-India Radio, said the enrichment project was part of BARC's efforts to develop expertise in all aspects of nuclear energy and that there was no plan to set up a bigger plant.

India has no need for enriched uranium as its Candu-type reactors are based on natural uranium fuel and heavy-water moderator. Only the two US-built light water reactors (LWR) at Tarapur at present depend on U-235 fuel imported from France. Under an agreement with the United States, US permission is required for operating the reactors with India's own enriched uranium fuel. According to Iyengar, the enriched uranium will be required for nuclear physics experiments.

India is self-reliant in the total nuclear fuel cycle technology — from mining uranium and commissioning power plants to reprocessing spent fuel. Enrichment was one area where it lagged behind. Having developed the gas centrifuge technology, India is now on equal terms with Pakistan, but AEC sources have dismissed speculation that India's enrichment activity is re-

lated to a weapons programme.

US news reports that Pakistan is on the verge of building a bomb have undoubtedly caused alarm in India, but the government has said it will not take any hasty decisions about going nuclear. Replying to questions about the possible nuclear threat from India's neighbour, Mr Natwar Singh, Minister of State for External Affairs, told parliament that India "will keep its powder dry" while at the same time exploring all avenues for preventing a nuclear race in the subcontinent. India, which conducted a nuclear test at Pokhran 12 years ago, has consistently denied having a weapons programme.

Meanwhile, AEC has announced an ambitious plan to build 22 more nuclear power plants in the next 15 years to raise India's nuclear power capacity from the present 1,100 MW to 10,000 MW. Ten of the reactors will be 500-MW units (now in the design stage) and the rest will be similar to the 235-MW plants in operation in Madras and Rajasthan. The 15-year plan would cost \$14,000 million at 1983 prices. Because government resources are limited, it has been decided to bring nuclear power generation under a public limited company that will raise funds through the issue of bonds and shares. Although AEC is confident of launching the Nuclear Power Corporation, it remains to be seen whether the Indian public will invest in this programme after Chernobyl.

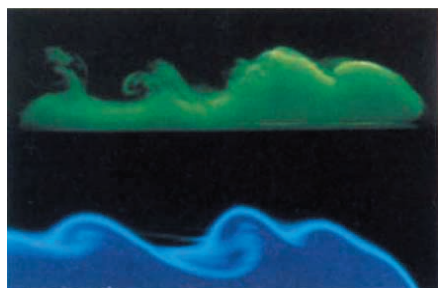
K.S. Jayaraman

Storm flow simulation by computer

Washington

A NUMERICAL simulation of thunderstorm outflows by Kelvin Droegemeier of the University of Oklahoma is being used to program a 727-passenger jet flight simulator at the Federal Aviation Administration's training academy in Oklahoma City. It is hoped that pilots can learn to recognize and avoid dangerous wind shear associated with outflows, which has caused several fatal crashes.

Droegemeier has produced a film of his simulations, which were run on a Cray X-MP supercomputer at Omnibus Simulation Inc. of Los Angeles. The blue image shows a still from the numerical model; the



green image shows a laboratory tank study in which the same phenomenon is modelled by releasing salt water into less dense fresh water. The simulation, which has about 40,000 grid points, reproduces clearly the features seen in tank studies 20 years ago. The turbulent eddies on top of the outflow in both images are corroborated by Doppler radar images of real outflows, and probably explain the seemingly random pressure fluctuations observed on the ground beneath. Production of Droegemeier's film was paid for by the National Science Foundation's Office of Advanced Scientific Computing. Tim Beardsley

AIDS

German survey's gloomy outlook

THREE-QUARTERS of those infected with the AIDS (acquired immune deficiency syndrome) virus will enter the final and fatal stages of the disease within seven years, according to a research team from the University of Frankfurt.

The predictions from the study, claimed to be the first long-term project in Europe, are based on "continuous observation" of 543 patients, largely the sexual partners of the first AIDS victims to die in Frankfurt, who have been attending outpatients' clinics since 1982. On the basis of the rate of deterioration of these patients — 377 proved positive in original antibody tests — and from computer forecasts, the researchers conclude that half the antibody carriers will have progressed to the final stage of AIDS within five years of first exposure and three-quarters within seven years.

The research team is led by Professor Eilke Brigitte Helm from the Zentrum der Inneren Medizin at the University of Frankfurt and the findings are published in the *Deutsche Medizinische Wochenschrift* (No. 32/32, 1986). The team placed all the patients in one of four categories, depending on the seriousness of the symp-

toms. They were those whose antibody tests proved positive but who had no obvious sign of the disease; those with a slight immunological weakness and showing symptoms of lymph-node swelling; those with pronounced immunological defects; and those with the full symptoms of the disease. The progress of the patients was monitored for three years.

At the end of that period, only 30 of the original antibody-positive patients were still healthy and, within a year, half of the patients with no symptoms moved into the next category. Within two years, half of those with slight or pronounced immunological defects were entering the final stages of the disease or suffering severe disintegration of their defence systems. With one exception, all those in the final stages of the disease died within two years.

According to the researchers, the belief that AIDS would develop only in a minority of cases infected with the virus is unfounded. The predicted size of that minority has increased with time from 5 or 10 per cent to at least 30 per cent in some cohorts of patients, but the German prediction of 75 per cent is unprecedented.

Bill Johnstone

Technology

Genentech buys out partnerships

Washington

GENENTECH Inc., one of the leaders in US biotechnology, has decided to reap an early harvest by announcing plans to buy out two of its research and development partnerships early. A limited partnership is a means by which outside investors support part of a company's research and development programme, sharing the profits if it succeeds and otherwise being able to set off the loss against taxes.

Genentech Clinical Partners Ltd (GCP) and Genentech Clinical Partners II (GCP II), have provided support and shouldered risk while Genentech's seedling products, including the human growth hormone Protropin, tissue plasminogen activator Activase and gamma interferon are being developed and tested clinically.

According to the original agreements, Genentech would not have bought out the partnerships until two years after the products from each of them had received approval from the US Food and Drug Administration (FDA). This would have been 1988 for GCP, as FDA approved Protropin this year, and a later date for GCP II, depending on FDA approval for the marketing of Activase.

Buying out the partnerships will give Genentech exclusive US rights for Protropin, Activase and gamma interferon, and all the associated profits. The partners will be offered 3,000 shares of Genentech common stock for every \$50,000 partner-

ship unit in GCP and 2,500 shares for every similar unit in GCP II. If the stock price on the date of the transactions remains at \$80, as expected, each \$50,000 investment in GCP would receive Genentech stock valued at \$240,000 and in GCP II stock valued at \$200,000.

By acting now, before the partnership agreements have expired, Genentech avoids buying completely developed products that would have to be capitalized and amortized. This development shows Genentech's faith in the products developed under the partnership agreements; on Wall Street, Genentech's stock rose by \$12.75 per share on the day after the proposed buyout was announced. The price is now about \$85.

Genentech is one of the first US biotechnology companies to become profitable. It recorded a profit of \$5.6 million (\$0.18 per share) in 1985, and its financial report to the end of September this year shows a net income of \$8.2 million, an earning for each share of \$0.23. Although the early promise of big profits for 1986 will be diluted by the one-off expense of the GCP II buyout, profits for 1986 are likely to be double those for 1985.

Genentech is also now planning a two-for-one stock split and intends, like many other leading US companies, to reincorporate in the state of Delaware to benefit from the favourable legal climate there.

Carol Ezzell

Student squeeze in West Germany

Hamburg

WHILE the universities in West Germany are still groaning under the burden of too many students — 1.3 million are sharing 850,000 places (see *Nature* 322, 197; 1986) — they are preparing to compete for students in the decade ahead. The number of school students is falling; the era of expansion is coming to an end.

A detailed analysis carried out at the centre for the study of regional development at Giessen and sponsored by the Deutsche Forschungsgemeinschaft (DFG) has shown that the new student generation will prefer big and old universities, so that the new and smaller institutions, particularly in Bavaria, may "have to fight for their survival".

The study is based on official university statistics, interviews with school leavers and figures from the Zentralstelle für die Vergabe von Studienplätzen (ZVS). The results are alarming for the cities concerned. Giessen, for example, with 17,500 students among its 70,000 inhabitants, will lose more than 4,000 students.

Other universities, many of which are at

present overcrowded, will not suffer so greatly. The analysis, which was carried out by Professor Ernst Giese, concluded that 22 out of 52 West German universities are "above average", in being chosen by more than 2 per cent of new undergraduates. The University of Munich attracts 8 per cent, followed by Cologne, Hamburg, Hannover, Münster, Munich (Technical University), Aachen, Bonn and Berlin (Free University) with between 5.49 and 3.5 per cent.

Old universities are preferred because of their traditions but also because of the recommendations of previous students. Universities such as Bremen, Frankfurt and Marburg, which were considered to have left-wing tendencies are well regarded now.

One of the ways in which the universities will compete is by setting up special study centres concentrating on specific subjects such as solid-state physics at Stuttgart, genetic engineering at Heidelberg and Würzburg, plant biotechnology at Hamburg and Cologne and information technology at Karlsruhe.

Jürgen Neffe

Chinese-US collaboration

NSF refuses fees

Washington

THE National Science Foundation (NSF) is digging in its heels against what it sees as unreasonable charges levied by the Peoples' Republic of China for US field trips in the country. Officials say the Chinese policy of charging for field trips is contrary to the spirit, if not the letter, of government agreements on scientific exchanges, and have given notice that NSF will not pay for a return visit next year by a US team taking glaciological samples in the Quilian mountains in Tibet.

The US team, headed by Lonnie Thompson of Ohio State University, spent 28 days last summer in a remote glacial area 1,000 km west of Lanchow, working with scientists from the Lanchow Institute of Glaciology and Geocryology. According to NSF's Eugene Bierly, China at first asked for \$100,000 to defray expedition expenses, but after negotiations, the figure was reduced to \$80,000. Thompson plans to return next summer to bring back ice-core samples, but NSF says it will not meet field-trip expenses again on the grounds that they violate the "receiving side pays" convention that underlies the US-Chinese agreements.

Bierly said there would be no question of charging Chinese scientists for field trips in the United States, and expressed the fear that budget cuts at some Chinese research institutes might tempt them to make deals with scientists from other countries, perhaps allowing priority of access to those who pay well. Bierly said NSF is determined to "hold the line".

Others report similar difficulties with fees for access to China. Craig Black of the Los Angeles Museum of Natural History said last week at a meeting of the National Science Board that charges for biological field trips were commonplace and that it was "almost becoming competitive". Officials at the meeting were alarmed by reports of a Japanese-Chinese agreement on glaciological research according to which, it is said, Japan will pay \$1,000,000 and donate four expedition vehicles in return for access to a large tract of China. NSF officials say the foundation would be unable to offer funds on that scale.

Thompson disputes NSF's figure of \$80,000 for the fee for last year's Quilian expedition, saying it was only \$40,000. The fee was, he says, reasonable for the level of logistical support provided, and that it "would be a good deal in any other country". NSF meets expedition costs in countries with which there are not collaborative agreements. But Thompson says he has agreed with the National Geographic Society that it will pay the expedition costs next year, so NSF's stand should not imperil his venture.

Tim Beardsley

Rhine pollution

Death of Europe's sewer?

In the French city of Strasbourg, on a tributary of the Rhine just 100 km downstream from Basel, feelings are running high, as they are all along this long river. Demonstrators gathered in Strasbourg and other riverside towns on Sunday carrying placards with messages such as "Poison ou poisson: il faut choisir" ("Poison or fish: we must choose"). Some fish have survived, but are polluted, and fishing has been banned for six months along the upper Rhine and its canals.

The Rhine seems 'dead' for 100–200 km downstream of Basel in northern Switzerland after the fire in the Sandoz chemical plant led to the discharge into the river of up to 32 pesticides, fungicides and other agricultural chemicals. One of the most important, 'disulfeton', is said to be twice as toxic as potassium cyanide; the average dose to kill 50 per cent of rats (the LD₅₀ criterion) is about 5 mg kg⁻¹ compared with 10 mg kg⁻¹ for potassium cyanide.

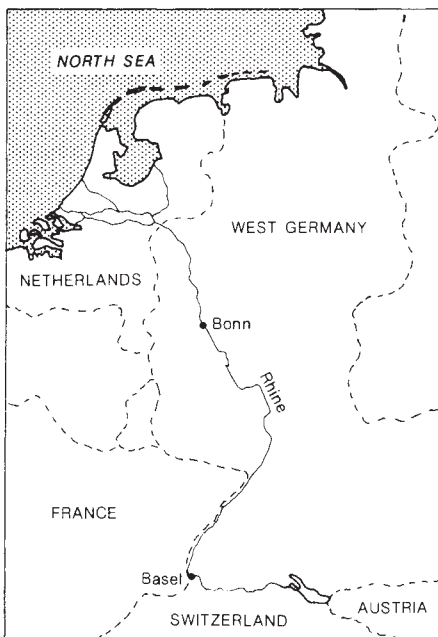
These facts emerged after the owners of the factory made public a list of the 1,000 tonnes of chemicals stored on the Basel site. Disulfeton has been detected as the principal pollutant from the accident as far as the Dutch border, 600 km downstream from Basel; but by this point the disulfeton had been greatly diluted to around 2 mg l⁻¹ of Rhine water, according to informed sources.

Sandoz and the Swiss Federation have admitted liability for the accident, which occurred after firemen had flooded Sandoz's Basel factory with water during a fire, to the extent that a dyke was breached and a toxic 'soup' of some 10–30 tonnes of agrochemicals was released into the Rhine. Mercury-based fungicides such as tillex as well as disulfeton and other organophosphorus compounds were in the mixture, which killed fish, normally resistant eels and insects. The effects on plant life are as yet unknown, and although little organo-mercury was detected downstream, there are fears that such chemicals have been deposited as sediment close to the release point, where they may be slowly released to the food chain for some time.

Thus, according to European environmentalists, the damage has been "enormous". According to the European Environmental Bureau, an organization that acts in Brussels for a number of European environmental groups, the Rhine was beginning to recover until this accident. Industry had been reducing its emissions under a combination of voluntary and statutory controls, to the point where fishermen were beginning to look for salmon at Basel. But this accident released as much pollution in a couple of hours as the Rhine normally receives in a year. "It's set the

clean-up back a decade", a bureau spokesman said.

Meanwhile, people are despondent about the months ahead. The proprietor of one fishing tackle shop said she had had no customers for 3–4 days and expected none now until the spring. Did she



expect any compensation from the authorities? She gave a resigned shrug of the shoulders.

Nevertheless, there is no doubt that the accident has set the European environmental movement and the legislators a good few steps forward, just as the Chernobyl accident gave a boost to the anti-nuclear lobby. Last week, the European Commission was pointing out how many countries have failed to put into effect the 'Seveso directive' of 1982, in which European Community countries agreed (subject to the passage of relevant national legislation) to impose controls on chemical factories. This followed an accident at Seveso in Italy in which a large area was contaminated by dioxin.

But four years after the directive, the Commission is satisfied only with legislation enacted by Denmark, France and Britain. The Netherlands, one of the loudest complainants against the Sandoz accident, has failed to apply the Seveso directive's principles on worker protection; West Germany has problems in the application of rules on the storage of dangerous substances; Greece's legislation has been too vague and "inadequate"; and Italy, most affected by Seveso, has not applied one of the directive's most crucial elements, a requirement to be placed on companies regularly to inform the authorities and their workers of the processes taking place

in a factory, of the chemicals stored and of emergency plans.

The Commission is now planning action against defaulting countries, but to begin with is only taking on a minnow — Luxembourg — in the European Court; with the rest, it has satisfied itself with sending stern diplomatic notes.

Meanwhile, countries belonging to the governmental International Commission for the Protection of the Rhine met in Zurich last week, heard contrite Swiss apologies and offers of damages (to a scale as yet undetermined), and set in motion the investigation of measures that would considerably tighten up controls and safeguards on Rhine industries. Switzerland itself will consider adopting national legislation along the lines of the Seveso directive of the European Communities — although as a non-member it has no obligation to do so — and there is now great pressure for Switzerland to sign an international convention that would oblige Swiss companies to inform their national neighbours of their activities, safety procedures and contingency plans.

The members of the International Commission have also decided in principle to apply to the Rhine the principles of the European convention on the reporting of oil releases at sea (Switzerland was very late in informing its downstream neighbours of the Sandoz release) and to set up an insurance system equivalent to TOVALOP, the tanker owners' voluntary agreement on liability for oil pollution. Any new conventions of this kind would of course apply to all forms of pollutant release and not just to oil. **Robert Walgate**

• Chemical accidents and spills are not confined to Western Europe. During the past few weeks, several incidents have been reported from the Comecon bloc.

On 1 November, at the chlorine and PVC works in Devnya (Bulgaria), there was an accident in which several people were killed, which destroyed the control rooms of two sections of the plant and also the finished products store, and which was apparently caused by rupture by an inadequately maintained pipeline.

On 11 November, in the Buna chemical works in East Germany, a hot-condensate condenser was destroyed by excess pressure, with four workers seriously injured and "considerable damage to property". And on 12 November, Prague radio reported that an accident in the Ostrava sewerage system had led to the discharge of heating oil into the rivers Lucina and Ostravica, and then into the Oder. The damage, the Chechs said, was expected to be long-term because of contamination of the river banks, but "drinking water has not been affected". According to the East German Ministry for Environmental Protection and Water Management, the spillage was not expected to affect the East German reaches of the Oder. **Vera Rich**

Genetic manipulation

Living outside regulation

Washington

A GENETICALLY altered strain of yeast designed to produce a low-calorie beer will soon be churning out trial brews at the Brewery Research Foundation (BRF) in Surrey, England. BioTechnica International of Cambridge, Massachusetts, chose to test its new creation outside the United States in part to postpone facing the regulatory process awaiting it in this country.

BioTechnica has so far produced only 10 millilitres of beer in the laboratory to test the process, and is prohibited from producing more than 10 litres by the Cambridge Biohazards Committee. To do testing on a larger scale, BioTechnica sought out BRF in England because there is no equivalent brewing test facility in the United States. But Roger Yocum, director of BioTechnica's yeast research division, makes no bones about the fact that testing the new yeast will probably be much easier in England.

BioTechnica International has been working for more than three years to produce a strain of *Saccharomyces uvarum* (also called *S. carlsbergensis* or brewers' yeast) containing a gene from the mould *Aspergillus niger* encoding for a glucoamylase enzyme. So-called lite (for 'light') beers are less fattening because they contain less starch than standard brews. By introducing a glucoamylase into brewing yeast, the beer can be produced faster and without other additives to remove the starch.

John Hammond, a biochemist at BRF, says that so far all work on the yeast has involved less than 5 litres, and all brews have been pasteurized to kill the yeast. But an unpasteurized beer, such as is provided to bars and restaurants, could contain live yeast, raising the issue of environmental release of what is an intergeneric organism. Hammond says BRF will consult with the Advisory Committee on Genetic Manipulation and other relevant governmental agencies before proceeding with any large-scale work with BioTechnica's yeast. But Yocum hopes that any US decision whether to permit environmental release of the new yeast will be clarified by BRF's testing.

David Kingsbury, assistant director of the National Science Foundation and one of the designers of the new government regulatory framework for biotechnology, says that even though the new strain of *S. uvarum* is probably safe and not an environmental hazard, it would come under the highest level of federal scrutiny because it is an intergeneric organism. Kingsbury says this shows that the new federal guidelines are if anything too strict, rather than too lax as some critics have charged. In the case of the new yeast,

the Food and Drug Administration (FDA) would probably have initial regulatory authority. Henry Miller, special assistant to the commissioner of the Food and Drug Administration for biotechnology, says the fact that the yeast is an intergeneric organism would have no particular relevance to FDA's regulatory process. FDA's charge is to determine whether it is a food additive, and whether it transforms a product that is already "generally recognized as safe".

Testing outside the United States is hardly a panacea, as the Wistar Institute in Philadelphia, Pennsylvania learned recently. Wistar and the Pan American Health Organization (PAHO) had been conducting a field test of a rabies vaccine using a modified vaccinia virus containing the sequence for an immunogenic protein from the rabies virus. Twenty cows were injected with the modified virus — ten subcutaneously and ten intramuscularly — and housed with twenty unvaccinated animals.

In September, news of the experiment appeared in the Argentine press, stirring a political controversy. That PAHO had not informed the Argentine government that the field trial involved a recombinant DNA organism made matters worse. PAHO has been conducting trials on rabies vaccines at a field station in Azul, 180 miles south of Buenos Aires, and it is unclear whether they had a legal obligation to inform Argentine officials.

Privately, a PAHO official concedes it would have been politic, given the politically sensitive nature of recombinant research, to inform Argentine authorities. Rabies is a significant problem in Argentina, with millions of dollars in cattle lost annually from the disease. Preliminary results from the Azul experiment show that the inoculated animals did show an immune response to rabies, while the uninoculated animals and the four human caretakers did not.

The furor in Argentina is ironic as Wistar associate director Warren Cheston says an Argentine government research laboratory contacted the Wistar Institute in 1984 to ask about conducting field trials with the vaccine. At that time Wistar declined, saying additional laboratory testing was needed.

A team from the Oregon State University in Corvallis had better success testing its modified vaccinia virus. The university announced last week that a field trial for a vaccine for Sindbis virus had been carried out over the past year in New Zealand. Calves, sheep and chickens were vaccinated in the study. None developed disease or transmitted the virus to unvaccinated animals.

Joseph Palca

India

Move to stop sex-test abortion

Bangalore

CONCERNED over the steady increase in the incidence of female feticide, the government of India is planning legislation to restrict the use of techniques for determining the sex of a fetus. The central health and family welfare ministry in New Delhi is seeking advice from medical experts, social workers, lawyers and women's groups in formulating the legislation. The aim is a comprehensive bill that would drastically limit the practice of amniocentesis and other emerging techniques of prenatal sex-detection.

As it is, the Medical Termination of Pregnancy (MTP) Act of 1971, which stipulates that an abortion can be performed only if there is risk to the life of the pregnant woman or can cause grave injury to her mental or physical well-being, has been ignored by some of the medical fraternity.

For many months, women's groups in Bombay and New Delhi have not only been campaigning against sex-linked abortion but have also been demanding that the authorities provide effective legislation to stop the abuse and to prevent the spread of clinics that exploit the legal loophole.

Mr R.P. Ravindra, a social worker in Bombay, points out that amniocentesis as a tool for sex-determination is almost always followed by abortion if the fetus is female. He says that a study of 8,000 cases of abortion showed that 7,997 of the fetuses were female.

Altering the sex ratio (at present 935 females to 1,000 males) through the abortion of female fetuses could adversely affect the social balance. India has one of the highest maternal mortality rates in the world. Moreover, about 70 per cent of women in India suffer from anaemia and malnutrition, both of which are likely to be aggravated by repeated abortions after sex-determination.

Supporters of the tests claim that the number of unwanted and neglected female children will be reduced and that a scarcity of women will raise the status of women in general. Others believe that the result will be polyandry and an increased incidence of rape and prostitution.

Not much is known about the numbers of sex-test clinics or their operation. Many are run under the guise of maternity homes, clinical laboratories and family health centres. Many are in the industrially and agriculturally advanced states of Maharashtra, Punjab, Haryana and Delhi. There are 248 clinics in Maharashtra according to the state health department.

Radhakrishna Rao

Japan

Body light points to health

Tokyo

THE Research and Development Council (RDC) of the Japanese Science and Technology Agency has sufficient confidence in Professor Humio Inaba of Tohoku University to invest about Y15,000 (£70 million) in his research. For Inaba believes that the light people emit may reflect the state of their health.

Inaba's team has found that the blood of patients suffering from diseases such as cancer, diabetes and jaundice gives off much more light as that of healthy people. Smokers' blood "glows" twice as brightly as that of non-smokers, but returns to normal after a day of not smoking.

Although the mechanism of photo-emission in each of these cases has still to be determined, peroxidation of lipids through the generation of activated oxygen and free radicals is believed to be involved. Peroxidal lipid reactions are suspected agents in diabetes, liver and lung diseases, arteriosclerosis, ageing and cancer. Inaba believes biophotons will eventually be used for diagnostic and clinical purposes.

The biophoton project is one of the latest additions to Japan's ERATO programme, which was set up in 1981 in a radical attempt to break free from Japan's conservative research system.

ERATO projects run for five years and, with the sole exception of the project leader, the 20 or so participants are all young researchers (under 35) drawn equally from universities, government research institutes and industry. Two or three foreigners are usually included in the team and

the research is carried out in independent laboratories uncontaminated by the usual environment of Japan's academic and government research laboratories. Once the theme of the project is chosen, the teams are left to find their own path, sometimes with bizarre consequences.

Inaba, who is attached to Tohoku University's Research Institute of Electrical Communication, became involved in maser research in the late 1950s, during an unsuccessful bid to get a large radiotelescope built in the Tohoku area north of Tokyo. In the 1960s, he turned to the development of high-sensitivity detectors for lasers, and by the end of the decade had devised a laser system for the detection of air pollution using highly sensitive photomultipliers to pick up the scattered light (*Nature* 224, 57; 1969).

Using Inaba's equipment, biochemists trying to detect the low levels of light ex-

pected to be given off during biochemical reactions involving peroxidation were able to detect ultraweak light emissions (of the order of tens of photons per second) from tumours and regenerating livers in mice and rats. But dark noise prevented determination of the spectrum of the emitted light.

Inaba's team, with support from RDC, went on to develop a luminescence analyser in collaboration with Tohoku Electronic Industrial Co. Ltd that can detect less than one photon per second and can handle samples up to the size of a rabbit.

The project will aim at improving this technology to allow measurement of weak biophoton emissions and to bring spatial resolution of the sub-micron level.

The technology has already found applications in the salad oil, rubber, paint and powdered milk industries, where the degree of oxidative deterioration of the product can be checked in a few minutes by monitoring ultraweak chemiluminescence. But medical applications seem to hold most promise. **David Swinbanks**

Nuclear safety

NRC investigates Seabrook

Boston

THE alleged misuse of drugs and alcohol by construction workers may have compromised the safety of the Seabrook Nuclear Power Plant in Seabrook, New Hampshire, which is due to begin low-power testing at the end of this year. The Nuclear Regulatory Commission (NRC) is now investigating these allegations.

The outcome of the inquiry is likely to affect both the debate about mandatory drug testing and that about existing controls on the safety of nuclear power plants. Drug tests are opposed by those who consider them an infringement of personal liberty.

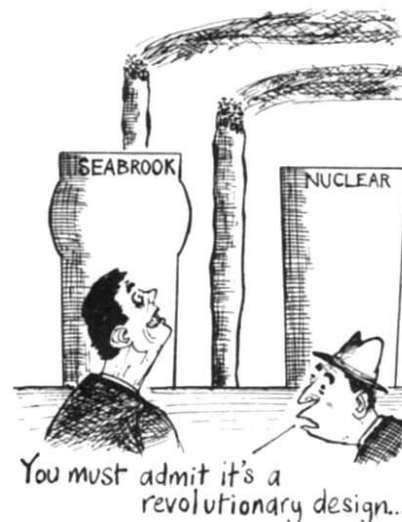
Several groups, including the energy conservation and power subcommittee of the House of Representatives headed by Congressman Edward J. Markey (Democrat, Massachusetts), have gathered allegations from police, hospital personnel and court employees as well as from the construction workers themselves. The alleged abuses are said to have transformed the town during the ten years of construction.

Former plant workers called Seabrook "a drug addicts' haven". Police have reported finding steel workers "drunk, coming and going" and, according to Dr Gary Lamphere of Exeter Hospital, "it was unusual enough in the early 1980s to see a Seabrook patient without alcohol on his breath to remark on it".

The issue is not whether workers constructing nonvital areas of the nuclear power plant were drunk or high on drugs, says Thomas Elsasser of NRC's Pennsyl-

vania office, but whether "systems essential to safety" have been affected. NRC will judge Seabrook's safety by inspecting documents, examining systems on a sampling basis and monitoring containment and fuel-load tests.

So far, Seabrook has passed all tests:



21,000 hours of inspection by NRC, almost twice that for any other plant in the region, have uncovered no structural reason to believe the plant is unsafe.

But testing after construction is considered unsatisfactory because some parts of the plant are no longer accessible for inspection. NRC may face other obstacles: it is alleged that some quality control inspections were not completed properly nor documented correctly.

Elizabeth Collins

SERC annual report

THE UK Science and Engineering Research Council (SERC) has asked the government, through the Advisory Board for the Research Councils, for an extra £20 million to make up the shortfall in its budget for 1987-88 due to unfavourable fluctuations of the pound against European currencies. The disclosure was made last week on the publication of SERC's annual report.

Professor E. W. J. Mitchell, chairman of the council, in the introduction to the report, claimed: "Overall the council's task remains the achievement of a sensible balance between curiosity-driven basic research, the seed corn of the future, and the strategic work where application might be anticipated in the medium term. This task is made more difficult within a severely constrained budget, further exacerbated by the wayward behaviour of the exchange rates which in 1985-86 were invariably against us." **Bill Johnstone**

Queensland and Creationism

SIR—In response to inaccurate statements made by Tony Thulborn (*Nature* 315, 89; 1985), I am writing to clarify my department's policy on the teaching of the theory of evolution in Queensland State secondary schools.

The Queensland Education Department policy on the teaching of the theory of evolution requires that it be taught (1) as a theory, not as fact; (2) in a balanced manner.

In providing balance, teachers must acknowledge that alternative theories to evolution exist. Some may be supported by scientific evidence; others may not be scientific in nature, but nevertheless based on beliefs deeply held by a significant proportion of the community.

Beyond this, I believe that the question of what constitutes "balance" should be left to the professional judgement of each and every science teacher. I hasten to add that I would be most concerned if a teacher taught the theory of evolution in such a way as to assail the religious beliefs of students. I would be equally concerned if the intent of a science syllabus, in terms of both "content" and "minimum time", were not adhered to.

LIN POWELL
(Minister for Education)

Treasury Building,
Queen Street, Brisbane, 4000,
Australia

Sexism and conservation

SIR—The advertisement on the back cover of the 16 October issue of *Nature* shows a young woman clad in only a shell necklace. While past correspondence has discussed the issue of sexism in advertisements in your journal, we want to point out that many gastropods are fished and killed for the explicit purpose of making jewellery because local entrepreneurs have exhausted the dead shells that are washed up on beaches. In certain countries, especially those with tropical shorelines, this practice has reached epidemic proportions and constitutes a major threat to the resident molluscs¹. Display of material that links (live) girls to (dead) prosobranchs can only damage efforts to conserve sea-shore fauna.

T. M. CARE
Sub-Department of Animal Behaviour,
University of Cambridge,
Madingley, Cambridge CB3 8AA, UK

W. K. LINDSAY
Department of Applied Biology,
University of Cambridge,
Pembroke Street,
Cambridge CB2 3DX, UK

1. Kendall, B. *Swara Magazine* 8 (1), 8-11 (1986).

Literary pedantry

SIR—In "What is the scientific literature" (*Nature* 322, 681; 1986), John Maddox appears to be confusing the medium with the message. Although I, too, decry the poor writing that characterizes many scientific publications, this alone does not drum papers out of the ranks of 'literature'. On the contrary, any scientific paper meets all standard dictionary defini-

tions of 'literature', with the possible exception of the definition that is characterized as *Archaic*, namely, for example, "literary culture" (*Dictionary of the English Language*. Random House, New York, 1968). Maddox is barking up an antique tree.

MELVIN BLECHER
Department of Biochemistry,
Georgetown University,
Washington, DC 20007, USA

Problems of Soviet refusniks

SIR—We beg the courtesy of your columns to address a letter, signed by us and seven of our colleagues, to members of the World Federation of Scientific Workers (WFSW) and its president, Professor Jean-Marie Legay.

We had hoped to hand this letter personally to Professor Legay when he was in Moscow this summer for the General Assembly of WFSW. Indeed, when we telephoned him there he courteously agreed to meet the two of us at his hotel on the morning of 24 July, before the opening ceremony of the assembly. On arrival at the hotel, however, we were denied entry and had to telephone Professor Legay again to explain our difficulty, which he promised to try to deal with. Instead, after some delay, we were forced out of the hotel and away from the entrance area by a group of plain-clothes men who brutally

informed us that we had "no reason to meet Legay". Subsequent telephone conversations established Professor Legay's inability to arrange a meeting with us during the remainder of his stay in Moscow.

The text of the letter appears below. We do not wish to embarrass Professor Legay and WFSW. Indeed, what happened on 24 July was as great a discourtesy to them, in being denied a visit by their invited guests, as it was an affront to us. Instead, what we wish to do is give WFSW, committed as it is to professional solidarity among scientists, the opportunity publicly to press for our professional rights and thus enhance its prestige as an organization dedicated to humane principles.

ALEXANDER IOFFE
BORIS KLOTZ
Moscow, USSR

To members of the World Federation of Scientific Workers

Dear Colleagues,

We believe that the WFSW, being an influential representative of the international scientific community, cannot ignore the plight of hundreds of scientists living in circumstances that leave them little chance to survive and preserve the ability to work in science.

We mean those who by one or another reason, national, religious, professional declared the intention to leave the USSR for Israel but was not granted the possibility to do so. Many continue their efforts for 10, 15 and even more years being practically deprived of any kind of scientific communication and, very often, of any real possibility to continue their professional activity. Many lost their jobs, a few even lost freedom. And all live under permanent psychological pressure, uncertain and uninformed about their future and the future of their families.

Meanwhile, we believe that the problem we are talking about is not among the most difficult-to-solve problems dividing the world. We believe that a just solution to the problem can easily be found if there is goodwill and understanding. We finally believe that, being found, such a solution would substantially contribute to the atmosphere of trust and cooperation between countries and societies.

The world scientific community has many times proved its determination and ability to defend human dignity and professional rights of scientists. We know of many activities undertaken by various scientific organizations and groups as well as by individual scientists on behalf of Jewish emigration in general or certain individuals whose circumstances were especially difficult. We call WFSW to join these efforts: the future of the world is inseparable from the future of individual human beings. Such a constructive demonstration of goodwill and professional solidarity would completely correspond to the humane principles proclaimed by the Federation.

Victor Fulmacht
Alexander Ioffe
Mikhail Kholmyansky
Boris Klotz
Yuli Kosharovskiy
Erlena Matlina
Vladislav Ryaboy
Igor Uspensky
Yosif Zaretzky

Moscow, 24 July, 1986

Mathematics for natural hierarchy

A new way of tackling the statistical problems of spin glasses relies on almost forgotten mathematics, but may have a bearing on problems in biology still unsolved.

How does it so often arise that the solution of newly defined problems hangs on the use of mathematical techniques independently developed, often decades earlier, for quite different purposes or (mathematicians being what they are) for no purpose at all? That question is the starting point for a delightful article in *Reviews of Modern Physics* for July this year in which the concept of "ultrametricity" is offered as an illustration of the principle that essential tools will usually be found lying discarded in the attics of the mathematicians. It goes without saying that the real world is not always as bountiful.

Part of the charm of this winsome piece is that its authors, R. Rammal (from the CNRS low-temperature institute at Grenoble), G. Toulouse (from the École Supérieure de Physique et Chimie Industrielles in Paris) and M.A. Virasoro (of the physics department of the University of Rome), confess that they themselves have only recently stumbled on the notion of ultrametricity, in an article on spin glasses in another physics journal. Like new converts in most fields, they have become enthusiastic proselytizers for the cause (*Rev. Mod. Phys.* **58**, 765; 1986).

Historically, so far as Rammal *et al.* have been able to tell, ultrametricity goes back to 1897, when the German mathematician Kurt Hensel published a recondite piece of algebraic number theory showing, among other things, that an arbitrary integer can always be represented as a sum of ascending powers of some other arbitrary integer, with each term in the series multiplied by a coefficient which may be zero but which is, in any case, less than the base for the ascending power series.

Nowadays, the most familiar illustration of the truth of Hensel's principle is the practice well known to computer buffs of representing decimal integers in binary form, which is merely a way of representing an arbitrary integer as a sum of powers of 2 in which sum the coefficients may be either 0 or 1. The more general statement is that the base of daily arithmetic may be changed from 10 to any other arbitrary base. A further generalization is that rational numbers (with finite, or non-recurring, decimal representation) may always be represented by the sum of powers of some arbitrarily chosen base in which some negative as well as positive powers are included. Rammal *et al.* say that Hensel's work was done in complete isolation, but its practical importance

now cannot be gainsaid.

The other mathematical innovation from which ultrametricity has sprung is the formal definition of metric spaces at the turn of the century. For a 'space', or a set of points, to be metric, there must be some function called distance of any pair of points which satisfies the condition that the value of the function is zero when the two points are the same, and that the value is unaltered when any pair of points is interchanged.

Euclidean verities come to mind, together with the euclidean principle that the sum of the distances between one point and a second and between the second and third is greater than the direct distance between the first and the third. This is just another way of saying that, in the euclidean world, the sum of the lengths of two sides of a triangle (say AB plus BC) is greater than the length of the third side (AC). Ultrametricity requires a more stringent principle, that AC should be less than whichever is the greater of AB and BC .

Such a metric world is, of course, an odd world. Rammal *et al.* point to some formal mathematics in the 1940s showing that, with definitions of metric distance satisfying these conditions, any point inside a sphere (defined as the region distant less than some fixed amount from a fixed point) will be itself at the centre of the sphere, and that the only possible triangles will be either equilateral or isosceles (in which case the base must be shorter than the two equal sides). Hensel's work is relevant because it provided one of the first illustrations of how such a metric might be defined.

What relevance can such an odd way of defining distance have to the problems of the real world? Rammal *et al.* start with a phylogenetic tree such as any cladist might construct. The tree implies a hierarchy which, once drawn, implies metric distances between the members at the lowest level, most easily thought of as the length of time since any pair of them diverged from a common ancestor. Then, in the comparison of the distances between three arbitrary extant species, either one pair has more recently evolved from a common ancestor than that evolved from some more distant common ancestor (in which case the triangle is isosceles) or no two of the three species have a common ancestor earlier than the putative common ancestor (in which case the triangle of the

distances between them is equilateral in the ultrametric sense).

From this point, it is an easy step to the application of the ultrametric concept to all kinds of intriguing problems in data processing, especially to the use of tree-sorting algorithms. But Rammal and his colleagues are proudest of their novel application of ultrametricity to the problem of spin glasses, the structures (real enough) represented by three-dimensional lattices whose vertices are occupied by spins (think of them as little magnets) and whose energy is the sum of the magnetic interactions between all possible pairs of vertices in the lattice. The problem is that of an Ising lattice in which the simplifying assumption that only nearest neighbours interact is abandoned.

Experimentally, the spin glasses are peculiar materials, especially at low temperatures. Phase transitions occur to states in which the solid material appears to consist of small domains of magnetically ordered spins, but the degree of hysteresis associated with changing external circumstances (such as temperature) is unusual; the materials seem to have a capacity for long-term memory of their previous state. In the decade since the problem of the spin glasses was recognized as such, much has been done by analytical and numerical approximations to show that the low-temperature condition of such a material may be one of a great number of essentially similar configurations that are distinguishable from each other largely by their symmetry. What Rammal *et al.* are able to show is that these configurations are related to each other hierarchically, by means of an ultrametric metric.

That will be an interesting landmark, but will it be an important one? Much depends on whether the authors' speculations turn out to be realistic. They claim, for example, that their line of argument may have a bearing on the travelling salesman problem (how to travel to N cities most economically), the founding conundrum of linear programming. Naturally, they say there is a place for neurobiology in their scheme of things. And since the folding of protein molecules into biologically active materials is a hierarchical process, why not subsume that as well? If they are right, a great many people will be learning a great deal of mathematics they did not previously know existed.

John Maddox

Biological conservation

National park boundaries and ecological realities

David S. Wilcove and Robert M. May

WHEN the United States Congress established in law the boundaries of Yellowstone National Park in 1872, the main motivation was the geology of the area, not the biology. The legislators wanted to create a park of sufficient size to protect the remarkable geothermal features¹ and the boundaries were drawn more or less as a giant rectangle, measuring about 99 by 86 kilometres, and containing well over 900,000 hectares of land. Almost immediately, some people began to wonder whether 900,000 hectares was sufficient to preserve the park's wildlife and natural features. Today, its population of grizzly bears is declining; each winter, elk move out of the park to a much reduced winter range; the state of Montana, fearing the spread of brucellosis to domestic livestock, wants to eliminate any bison that wander off the park; and geothermal energy development outside the park boundaries threatens to disrupt its geysers. A recent study² suggests at least seven out of eight national parks in western North America are too small to maintain minimum viable populations of several species found within them.

The lack of congruence between legal and ecological boundaries² is common to many parks. In Kenya, the seasonal movements of elephants and other game species take them well outside the boundaries of Amboseli National Park (39,200 hectares). When outside the park, the animals must compete with livestock and contend with poachers. In Costa Rica, three-wattle bellbirds (among other species) annually migrate from the protected montane forest of the Monteverde Reserve into unprotected lowlands. Where park boundaries cross ecological realities, it is usually because the park is too small relative to the space required by particular species, or too small relative to the scale of the natural and man-made disturbances that occur in and around it, such as fires, disease epidemics, insect outbreaks and seasonal or other variations in food supplies.

Territories

Large mammals, like the Yellowstone elk or the Amboseli elephants, need a lot of food and water and must travel far to get it. Their seasonal movements often encompass an area far greater than that of the parks in which they occur and illustrate the particular difficulties posed in the conservation of migratory species. Large predators, which usually occur at low densities, present different problems. As

some grizzly bears have home ranges of over 3,000 km², Yellowstone National Park by itself simply could not support a viable population of them. The park's grizzly bears are just a part of the larger population that includes those bears in the adjacent national forests.

In an attempt to assess this problem, Newmark² recently compared the legal and ecological boundaries of eight national parks in western North America. He defined the ecological ('biotic' in terminology) boundaries as the area encompassing the entire watershed of a park enlarged, if necessary, to encompass an area of sufficient size to maintain a 'minimum viable population' for the terrestrial non-flying species with the largest home range found within the legal boundaries. Setting the minimum viable population size at 500 — which is probably an underestimate — Newmark found that seven of the eight parks were too small by factors ranging

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Re-growing a national park — cow dung with guanacaste seedlings in Costa Rica (*Enterolobium cyclocarpum*). (D.H. Janzen.)

from 6 to 100. Even when the level was set as low as 50, several species of mammals could not be supported in the seven parks.

Newmark's findings are not unduly surprising, because none of these parks was set aside expressly to conserve particular mammals. Moreover, parks like Nairobi or Yellowstone (or even Central Park in New York City!) were originally created within a larger area of roughly similar 'wilderness' and so were effectively much larger when they were originally established. Many present-day problems have arisen gradually as settlement and civilization have grown up around such parks. But the magnitude by which these parks today fall short is alarming. Most of the threatened mammals in these western North American parks are top predators, and their disappearance from the parks could trigger further ecological imbalances.

The problem of the scale of natural disturbances is even thornier. How does

one go about determining the spatial 'requirements' of a forest fire, a pine beetle infestation, or an epidemic of rinderpest?

Until recently, most natural disturbances were unwelcome inside parks, as they were thought to harm wildlife and scar the landscape. Thus, Yellowstone National Park has a long tradition of suppressing forest fires, even though the lower elevation forests in the Yellowstone region may have burned as frequently as every 20–25 years before the arrival of European settlers³. In some areas, aspen stands, an important habitat for some species, are being replaced by conifers, apparently because they are not being rejuvenated by fire. Over the past two decades, park authorities have allowed naturally caused fires in remote areas to burn, provided they do not pose a threat to human life or property.

Disasters

On the other hand, for very rare and localized species or for very small parks, such natural disturbances can be catastrophic. Few conservationists would want to see a fire rage through the world's only stand of Virginia round-leaved birch, however natural such an event might be.

In terms of park design, a further consideration is whether to plan for an 'average' disturbance or a real disaster — the 'fire of the century' or the 'fifty-year flood'. Given the irreversible nature of most land development, it may be wisest to take a rather pessimistic view and to plan for the worst.

There are several ways to make park boundaries more compatible with ecological realities. The simplest solution in the long run is to make the designated parks big enough at the outset. Manu National Park in Peru includes an entire river drainage within its more than 1,600,000 hectares and may be one park that can handle almost any natural disturbance that befalls it. But opportunities to create parks of this magnitude are few and fast disappearing.

A second strategy is to link pre-existing parks, as long recommended by island biogeographers^{4,6}. Probably the finest example of this approach is the recent acquisition of La Zona Protectora in Costa Rica by a consortium of conservation organizations, including the Nature Conservancy, the National Park Service of Costa Rica, the National Parks Foundation of Costa Rica, the Organization of Tropical Studies and the World Wildlife Fund. La Zona Protectora is a 7,700-hectare stretch of forest 3–6 kilometres wide that links two important reserves in the country: the upper-elevation Braulio Carrillo National Park and the lower-elevation La Selva Biological Station. La Zona Protectora does more than merely increase the total amount of protected land; it attacks the problem of habitat

linking the two reserves, it has a much greater additive effect. The combined reserves can support larger populations of many species, and there is now a protected corridor for the approximately 35 species of birds that migrate from the uplands to the lowlands (plus other, less well-known, migratory mammals and insects).

Development

A third strategy is to allow development of the lands surrounding the parks, but to restrict it to those uses and practices that are compatible with conservation goals. Thus, many conservationists in the United States now speak of a greater Yellowstone ecosystem, referring to Yellowstone and Grand Teton national parks (total area 1,022,000 hectares), and parts of the nine national forests (total area 5,747,000 hectares) that encircle them. The management of these national forests is currently the subject of much controversy, as government officials and private citizens debate how activities like logging, mining and grazing should be regulated to avoid harming the flora and fauna of the region.

Another example of an arrangement which is good in principle, but where the actual working is more complicated, occurs in Kenya where the inner 518 km² of the Masai Mara Reserve are managed like a park, with no exploitation of the land permitted, whereas in the outer 1,294 km² the Masai tribesmen are allowed to graze their cattle⁷. The arrangement sounds fine, but encroachments on the core area are hard to curb. The rules in the Ngorongoro Reserve of Tanzania arguably work better: here the Masai can graze cattle at limited densities anywhere outside the crater, but may only take them into the crater for watering. Despite their problems, restrictive land-use arrangements such as these may be the only means to reconcile the conservation of wide-ranging species with human demands for the land.

Even where development has already harmed much of the land, it is possible to acquire small patches of pristine habitat, along with disturbed areas around it, and then to foster the outward growth of the park into the previously disturbed area. The current scheme to 'grow' a park in dry tropical forest in Costa Rica's Guanacaste Province is an example. The central protected block in Santa Rosa National Park is only about 60 km², but the eventual aim

is to enlarge the park to several hundred square kilometres by buying low-grade ranch, farm and timber land next to Santa Rosa and on the adjacent Volcan Orosi, and to let the park grow out into it. Only in this way is there likely to be enough area and moist upland refuges to buffer the severely threatened dry forest from normal environmental fluctuations, such as severe droughts and epidemics.

A campaign to buy this land, and create the larger Guanacaste National Park (for

a cost of around \$10 million) is now underway. If successful, it will demonstrate further that the existing network of parks does not have to be static as long as conservationists continue to come up with imaginative ways to enlarge and improve it. □

Robert M. May is Class of 1877 Professor of Zoology at Princeton University, Princeton, New Jersey 08544; David S. Wilcove is at The Wilderness Society, 1400 I Street NW, Washington DC 20005, USA.

Superstrings

In search of a smoking gun

Roberto Peccei

SUPERSTRINGS¹ have squarely split the elementary particle community into believers and sceptics. Superstring theories purport to describe a consistent quantum theory of gravity, unified with that of the other known forces, in which the nature of the matter constituents is also specified. No wonder, therefore, that superstring partisans abound. Unfortunately, all this unification occurs in 10-dimensional space-time, so that real physics can emerge from superstrings only as a result of the spontaneous compactification of the extra six spatial dimensions (see the *News and Views* article by Alvaro De Rújula in *Nature* **320**, 678; 1986). These must curl up into a little ball, torus or other compact shape of radius of the order of 10⁻³⁷ cm. Such scales will never be probed experimentally. Thus, say the sceptics, superstrings will forever remain figments in theorists' fevered imaginations. But a new breed of phenomenologists, superstring phenomenologists, is trying to bridge the gap between these two camps, by examining possible patterns left over from the process of compactification.

Although the energy scale of compactification at around 10¹⁹ GeV is enormous compared with that probed by present-day experiments (around 10² GeV), if superstring theories are correct, not very much is supposed to happen in between. Thus, one can hope to extrapolate smoothly through this energy desert, something which was attempted in earlier work on unifying the strong and electroweak forces into a grand unified theory². Also in superstrings, the usual forces are unified in a large symmetry structure before compactification. For the favoured heterotic superstring³ this is the group E₈ × E₈, where E₈ is the largest exceptional Lie group. In the process of compactification, however, this large symmetry breaks down to a smaller group. Clearly, for these theories to be sensible, this surviving symmetry group should be that connected with the known low-energy forces, or a small extension thereof.

There is another aspect of compactification which is crucial for four-dimensional physics. Superstrings have supersymmetry, a mathematical property which connects matter particles (half-integer spin fermions such as quarks and leptons like the electron) with the 'fields' that cause forces amongst them (integer-spin gauge bosons such as the photon, causing the electromagnetic force). The gauge bosons are associated with the symmetry group structure of the strings, and to each such boson supersymmetry assigns a half-integer spin partner (a fermion). On compactification, only some of these fermion partners survive. The kind and number of those which do are entirely fixed by the geometry of the compact space. If superstrings are to connect to reality, the surviving states must match fermions we see in four dimensions — quarks and leptons — with perhaps a few extra states yet to be observed.

So far there is no proof that the six unwanted spatial dimensions in superstring theories really can curl up, following the string equations of motion, into a space of a given geometry. In fact, it is not clear that if compactification occurs, there is a unique result. (There may be an infinity of possible worlds, of which ours has the charming property that some of its inhabitants may be able to understand it!) However, it is possible to adduce indirect theoretical evidence⁴ that, if this compactification takes place, the compact six-dimensional space is a so-called Calabi-Yau space^{5,6} or a generalization of these spaces, known as orbifolds⁷. Proceeding on this assumption, Candelas, Horowitz, Strominger and Witten⁴ derived two results which brought superstrings much closer to reality. What they found was that the emerging four-dimensional theory had a symmetry group E₆ × E₈ where E₆ is either E₆ or one of its subgroups, and that the concomitant four-dimensional fermions occurred only in replicas of a representation of E₆ containing 27 particles. This was good news for superstring fans,

1. Reese, R. *Greater Yellowstone: the National Park and Adjacent Wild Lands* (Montana Magazine, Helena, 1984).
2. Newmark, W. *Biol. Conserv.* **33**, 197 (1985).
3. Norse, E. et al. *Conserving Biological Diversity in Our National Forests* (The Wilderness Society, Washington DC, 1986).
4. Diamond, J. *Biol. Conserv.* **7**, 129 (1975).
5. Wilson, E. & Willis, E. in *Ecology and Evolution of Communities* (eds Cody, M. & Diamond, J.) (Harvard University Press, Cambridge, 1975).
6. Harris, L.D. *The Fragmented Forest* (University of Chicago Press, 1984).
7. Williams, J. *A Field Guide to the National Parks of East Africa* (Collins, London 1981).

as E_6 was already known as a nice grand unified theory group in which the low-energy forces fit well. Furthermore, the 15 quarks and leptons of each family fit perfectly within the 27-dimensional representation of E_6 . Matter and forces already known match precisely those emerging from this compactification, with a few extra forces and matter states thrown in for good measure.

The trouble is, one can have too much of a good thing. The simplest Calabi–Yau manifolds end up having enormous numbers (of the order of 100) of repetitions of the 27s of quarks and leptons. They also include forces that cause protons to decay very rapidly, totally at variance with observation. A systematic investigation of more favourable manifolds has been carried out by several groups^{8–10} in the past year, imposing reasonable boundary conditions to weed out uninteresting manifolds from the start. One of these conditions is that the resulting group, after compactification, should be so small that none of its gauge bosons can lead directly to rapid proton decay. In effect, this requirement necessitates that the E_6 group be broken at compactification. This obtains in manifolds of sufficiently complex topology so that it is possible for some of the E_6 gauge fields to become trapped in holes in the manifold, leading to a symmetry breakdown¹¹. Remarkably, precisely for these kinds of manifolds, the Euler characteristic, which details the number of matter replications, is also small. So two potential problems can be alleviated at the same time.

Two generic results emerge from these analyses^{8–10}. First, if all the symmetry breakdown occurs at the compactification scale, the surviving symmetry group contains the known forces plus at least one new force, whose associated gauge boson must have a mass comparable to that of the weak gauge bosons of around 100 GeV. Furthermore, the number of families is limited to three, and exotic states in the 27, besides quarks and leptons, survive at low energy. Second, if there are further stages of symmetry breakdown below the compactification scale, it is possible that the low-energy theory contains just the known forces. All exotic matter is, in this picture, superheavy and experimentally inaccessible; and up to four families of quarks and leptons are permitted.

Clearly the first case is more appealing phenomenologically, because it promises that both new forces and matter are there to be found, providing a 'smoking gun' to test the superstring hypothesis. However, this case is not guaranteed and is in itself not without difficulties. For instance, unless certain couplings of the exotic matter states are absent, protons will again decay rapidly. In this sense, the case with intermediate scale breaking is phenomenologically safer, but it is also almost devoid of

remnant signals. This is nicely exemplified by the work of Graham Ross and his collaborators¹² at Oxford, who studied in detail one of the few known three-family Calabi–Yau manifolds. At compactification, this manifold either retains the full E_6 symmetry or is broken down to $SU(3) \times SU(3) \times SU(3)$ or $SU(6) \times U(1)$. None of these patterns is nice. However, Ross and his collaborators¹² argue convincingly (although to my mind not totally compellingly) that a sequence of further break-downs take place which reduce the $SU(3) \times SU(3) \times SU(3)$ theory to precisely $SU(3) \times SU(2) \times U(1)$ — the gauge theory describing the strong and electroweak interactions. Not only are the forces standard at low energy but so is matter, with the exception of two weakly coupled neutral states. What remains as distinct predictions from the original theory are certain relationships between masses and mixing angles of quarks — predictions which are in fair agreement with experiment. Although this analysis is very nice, this is not much of a smoking gun to convince anyone who is not already an *aficionado* of the reality of superstrings.

In my opinion what will be necessary to sway superstring sceptics (among whom I count myself) is some hard evidence for the existence of superpartners to the low-energy excitations we know. Supersymmetry is, after all, a key feature of superstrings and, even after the Calabi–Yau compactification it remains a good symmetry of the four-dimensional theory. Although the matter fermions divorce themselves from the gauge bosons during compactification, they pick up in the process a set of spin zero partners so that, remarkably, a supersymmetry is retained⁴. Eventually, below the compactification scale this supersymmetry must itself break

down. How this happens is the least understood piece of the theory, although it is often suggested that the breakdown is triggered by physics associated with the other E_8 group which is unaffected by the compactification. At any rate, this breakdown of supersymmetry is necessary as it is the seed which allows all 'light' masses to be generated in the theory¹³, including the mass M_w of the weak gauge bosons. It is therefore natural, and in fact is required by the internal consistency of the theory, that the mass splitting among superpartners generated by this mechanism also be of the order of M_w .

Superpartners with masses of the order of 100 GeV are in the range of discovery for the forthcoming large accelerators at CERN near Geneva, DESY in Hamburg, Fermilab near Chicago and SLAC in Palo Alto. Their discovery, or that of some other debris of the superstring, would provide us with a first glimpse into extra dimensions. □

1. Green, M. *Nature* **314**, 409 (1985).
2. Georgi, H., Quinn, H. & Weinberg, S. *Phys. Rev. Lett.* **33**, 452 (1974).
3. Gross, D., Harvey, J., Martinec, E. & Rohm, R. *Phys. Rev. Lett.* **54**, 502 (1985).
4. Candelas, P., Horowitz, G., Strominger, A. & Witten, E. *Nucl. Phys.* **B258**, 46 (1985).
5. Calabi, E. in *Algebraic Geometry and Topology* (Princeton Univ. Press, 1957).
6. Yau, S.T. *Proc. natn. Acad. Sci. U.S.A.* **74**, 177 (1978).
7. Dixon, L., Harvey, J.A., Vafa, C. & Witten, E. *Nucl. Phys.* **B261**, 678 (1985).
8. Cecotti, S., Derendinger, J.P., Ferrara, S., Girardello, L. & Roncadelli, M. *Phys. Lett.* **156B**, 318 (1985).
9. Dine, M., Kaplunovsky, V., Nappi, C., Mangano, M. & Seiberg, N. *Nucl. Phys.* **B259**, 549 (1985).
10. Breit, J.D., Ovrut, B. & Segré, G. *Phys. Lett.* **158B**, 33 (1985).
11. Witten, E. *Nucl. Phys.* **B258**, 75 (1985).
12. Greene, B.R., Kirklín, K.H., Miron, P. J. & Ross, G. G. *Phys. Lett.* **B180**, 69 (1986).
13. Ibañez, L. *Phys. Lett.* **118B**, 73 (1982).

Roberto Peccei is at the Deutsches Elektronen Synchrotron (DESY), 85 Notkestrasse, D-2000 Hamburg 52, FRG.

Data bank for carbohydrate structures

THE need for an international data bank for carbohydrate structures was discussed at a recent workshop, sponsored by the US Department of Energy (Auburn, New York, 7–9 August 1986). Such a facility would provide primary structure information analogous to that obtained from the established amino-acid and nucleotide sequence data banks, which would allow the currently impossible systematic search for structural relationships between the 4,000–5,000 carbohydrate structures so far known.

The computer programs required to organize and store the carbohydrate structural data now exist but are more complex than those used for amino-acid sequences. The complexity results from the combinatorial effects of branching in the carbohydrate chain. Conference participants agreed that structural data would be collected by representatives of the different

sub-fields of carbohydrate research and organized centrally at the Complex Carbohydrate Research Center, University of Georgia, under the guidance of its director Dr Peter Albersheim.

The estimated costs for the first year of the project are \$200,000. A large contribution (40 per cent) will be made by the US Government and the remainder will come from the twelve other participating countries. Users will pay an annual subscription of \$100. This improved access to information will be welcomed by both academic and industrial (for example, pharmaceutical, recombinant DNA, oil) concerns.

An international steering committee was set up to coordinate the growth and funding of the project and the data bank should be in full use within a year. The conclusions of the workshop will be available in about two months from the US Department of Energy.

Sarah Hargreaves

Optical physics

Growing expectations from squeezed states of light

Dan Walls

THE past year has seen remarkable progress in the experimental generation of squeezed states of light, which previously were known only to theorists. Squeezed states have the property that the quantum fluctuations in one phase of the wave motion are less than those of the vacuum state of the electromagnetic field, although the fluctuations in the quadrature phase (90° out of phase with the other) compensate by being greater than in the vacuum. In this sense, in one phase squeezed states may beat the quantum limit imposed by the individual nature of the photons, the 'shot noise' of a light beam. There have recently been experimental reports of squeezed states of light generated by four-wave mixing in atomic vapour and optical fibres, and optical parametric oscillation. In the latter, a reduction in quantum fluctuations of more than a factor of two below the vacuum level has been achieved.

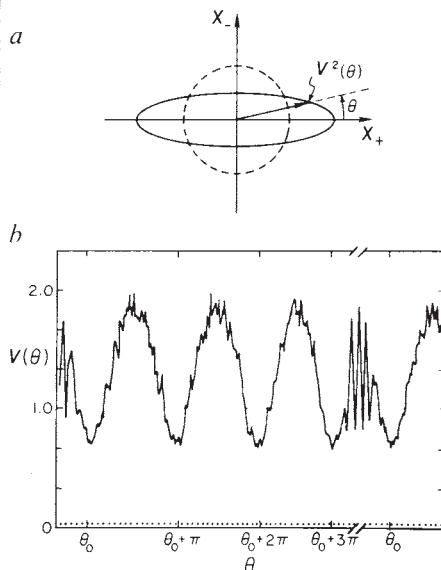
The interest in squeezed states of light arises from their possible application in gravitational wave detectors and optical gyroscopes, where they would enhance the sensitivity of interferometric measurements of phase. The reduced fluctuations in one phase offer the possibility of improving the signal-to-noise ratio in quantum-limited optical communications systems. New kinds of experiments in atomic spectroscopy are also feasible. (See my review of the general properties of squeezed states in *Nature* **306**, 141; 1983.)

Squeezed states are generated by nonlinear optical processes involving phase dependent amplification and deamplification of the phases in quadrature. Examples of such processes include four-wave mixing and optical parametric oscillation. The recent experiments examine different variations of these systems.

The first experimental results were reported one year ago by the group of the AT&T Bell Laboratories (Slusher, R.E., Hollberg, L.W., Yurke, B., Mertz, J.C. & Valley, J.F. *Phys. Rev. Lett.* **55**, 2409; 1985) in an experiment in a beam of atomic sodium. The beam is pumped with a coherent dye laser at 589 nm. To achieve a sufficient nonlinear interaction to squeeze the vacuum fluctuations, the laser frequency should be tuned near an atomic resonance. However, as spontaneous emission will destroy the squeezing, the pump laser is detuned from the atomic resonance by 300 atomic linewidths. The vacuum field in an optical cavity noncollinear with the pump laser is squeezed by the four-wave mixing process, and the

fields in cavity modes symmetrically displaced from the pump frequency (side-modes) by 80 atomic linewidths can be detected. This non-degenerate four-wave mixing process ensures the sidemodes are outside the spontaneous emission profile of the atom.

The quantum fluctuations in the output at the two cavity sidemodes are detected by beating with part of the laser which acts as a local oscillator. The local oscillator phase is varied with respect to the sidemodes to detect the quadrature with least fluctuations. The noise reduction observed in this experiment is 17 per cent below the vacuum level, in agreement with theoretical estimates for the parameters of the experiment. But theoretical calculations predict that squeezing should yield a factor of ten reduction in noise at higher pump powers. Another experiment using atomic sodium in a forward geometry in the Doppler broadened atomic vapour regime and no cavity has been performed at the Massachusetts Institute of Technology (Maeda, M.W., Kumar, P.



a, Phase plot of the uncertainties in the quadrature amplitudes of the electric field. The solid line represents the variance $V^2(\theta)$ of the field $X(\theta) = X_+ \cos \theta + X_- \sin \theta$ as a function of θ for a squeezed state; the dashed line is for vacuum fluctuations. *b*, Measurement of the phase dependence of the quantum fluctuations (noise power) in a squeezed state produced by degenerate parametric down conversion. The plot corresponds roughly to the quantity $V(\theta)$ as in *a*. The minima represent a 50 per cent reduction in noise power relative to the vacuum noise level. Note the ordinate is a linear scale in noise voltage. (Courtesy of H.J. Kimble.)

& Shapiro, J.H., personal communications). This experiment has to date achieved a squeezing level of 4 per cent below the vacuum. A further experiment with a sodium atomic beam investigates a new regime for which atomic and cavity timescales are comparable. Initial experiments at the University of Texas in Austin have obtained noise reductions of 13 per cent below the vacuum (Orozco, L.A., Raizen, M.G., Xiao, M. & Kimble, H.J., personal communication).

A four-wave mixing experiment but in a completely different type of material, an optical fibre, has been performed by a group at the IBM Laboratory in San José (Shelby, R.M., Levenson, M.D., Perlmutter, S.H., De Voe, R.G. & Walls, D.F. *Phys. Rev. Lett.* **57**, 691; 1986). Light from a krypton ion laser at 647 nm is transmitted through 114 m of glass fibre. Because the laser frequency is far away from any absorption level in the glass fibre, spontaneous emission is not a problem, but Brillouin scattering from acoustic phonons in the fibre introduces excess noise. The fibre is cooled to 4 K with liquid helium to reduce this scattering. The long interaction length in the fibre together with the high intensity of light focused in the fibre compensates for the rather weak optical nonlinearity of glass. Squeezing occurs in continuous pairs of sidemodes symmetrically displaced from the laser frequency. A heterodyne detection scheme is achieved by beating with the original strong laser beam which acts as a local oscillator and can be phase-shifted with respect to the sidebands. This system offers in principle squeezing over a broad bandwidth. However, mainly because of the excess noise from the Brillouin scattering the squeezing observed is limited to 12 per cent.

An experiment using a very different approach to the problem of squeezed state generation has been followed by a group at the University of Texas in Austin (Wu, L.A., Kimble, H.J., Hall, J.L., & Wu, H. *Phys. Rev. Lett.*, in the press) with spectacular success (see figure). The process of parametric down conversion, in which photons of frequency 2ω split to produce correlated pairs of photons at the sub-harmonic frequency ω , is used to generate 63 per cent squeezing.

The experiments are performed using a degenerate parametric oscillator consisting of a lithium niobate crystal inside a resonant cavity. The cavity is driven by light at the harmonic frequency (0.53 nm) by frequency doubling a Nd-Yag laser (1.06 nm). The down-converted light emerging from the cavity is combined with the original laser emission (at 1.06 nm) using a balanced homodyne detector to observe the squeezing. Although squeezing by more than a factor of two is observed, a quantitative comparison with theory suggests that in fact the intracavity

field of the optical parametric oscillator would produce more than ten times squeezing in the absence of losses in the current experiment. Kimble estimates that with improvements in detector efficiencies and cavity parameters the squeezing that can be observed will approach this milestone.

At this level many of the proposed ap-

plications of squeezed light look promising. Progress in this area has so far exceeded expectations and we can reasonably expect further exciting developments in the next year. □

Dan Walls is Professor of Physics at the University of Waikato, Private Bag, Hamilton, New Zealand.

Plant biochemistry

Discovery of the most abundant inhibitor in the world?

Christine Foyer

THE regulation of the activity of the primary enzyme of photosynthetic carbon assimilation, ribulose-1,5-bisphosphate carboxylase, has remained one of the most confusing and elusive areas of plant biochemistry but the recent discovery and identification of an endogenous nocturnal inhibitor, namely 2-carboxy-D-arabinitol-1-phosphate (Gutteridge, S. *et al.* *Nature* **324**, 274; 1986 and Berry, J. A. *et al.* *Proc. natn. Acad. Sci. U.S.A.*, in the press), should help resolve discrepancies in our understanding of the modulation of what must be one of the most important enzymes in the biosphere. Together with related observations of a light-dependant activation protein that facilitates optimum enzyme function at ambient levels of carbon dioxide (Salvucci, M.E., Portis, A.R. and Ogren, W.L. *Photosynth. Res.* **7**, 193-201; 1985), this work forms the basis for a new understanding of the overall regulation of carbon assimilation.

Ribulose-1,5-bisphosphate carboxylase was first identified in 1956 but it was not until the mid 1970s that sufficient activity could be achieved *in vitro* to account for the rates of photosynthesis found *in vivo*. Activation of the enzyme requires carbamate formation on a lysine residue of the large subunit followed by reaction with Mg^{2+} ions to form a catalytically active enzyme- CO_2 - Mg^{2+} complex but this knowledge provides only a glimpse of the functional capacity and regulation of the enzyme *in vivo*. Full activation of the enzyme *in vitro* requires high levels of carbon dioxide yet it occurs in leaves at ambient CO_2 concentrations. A further problem is that low concentrations of the enzyme activated with high CO_2 and high Mg^{2+} catalyse carboxylation showing distinctly biphasic kinetics with high rates of activity that decline after the first 15-30 seconds of the reaction to a lower steady rate.

A diurnal fluctuation in the specific activity of ribulose-1,5-bisphosphate carboxylase extracted from leaves has been observed in many plant species. Similarly, light activation and dark deactivation of

the enzyme have been demonstrated in leaves, leaf protoplasts and isolated chloroplasts. The activation state of ribulose-1,5-bisphosphate carboxylase in leaves can be correlated with irradiance in a manner that cannot be attributed to changes in stromal pH and Mg^{2+} . A dark-induced inhibitor of the enzyme may be responsible for this modulation. When leaves are illuminated, photosynthetic carbon fixation shows an induction phase (or lag period) of several minutes before maximum rates are attained. Previous results have suggested that light activation of ribulose-1,5-bisphosphate carboxylase is not an important factor in determining the duration of the induction period but the presence of 2-carboxy-D-arabinitol-1-phosphate adds an extra factor, at least in species where the inhibitor is abundant.

The presence of a dark-induced inhibitor of ribulose-1,5-bisphosphate carboxylase is at once intriguing and puzzling. The degree of activation of the enzyme found in darkness is highly variable and it is not clear why photosynthetic carbon assimilation requires that ribulose-1,5-

biophosphate carboxylase is relatively inactive in darkness. Several light-modulated enzymes are present in the pathway whose activity is negligible in the dark. These are activated on illumination long before photosynthesis overcomes the induction period. Now we find that there is another site of inhibition, and nature obviously requires a complete shutdown of the reductive pentose phosphate pathway in darkness.

2-carboxy-D-arabinitol-1-phosphate has so far only been identified in a limited number of species. There may, however be interspecific variation in the metabolism of the inhibitor and also differences in the conditions required for its retention during extraction procedures. We now need to know the biochemical pathways of synthesis and metabolism of this inhibitor and their regulation. Dr A. J. Keys (Rothamsted Experimental Station) has suggested that hamamelose bisphosphate could be a precursor of the inhibitor especially as this sugar-bisphosphate is made and rapidly turned over in the chloroplast. Synthesis and degradation may well be related to irradiance and therefore participate in the regulation of photosynthesis.

Ribulose-1,5-bisphosphate carboxylase has justly been called the most abundant protein in the world. As equimolar amounts of 2-carboxy-D-arabinitol-1-phosphate are required for optimum inhibition, this metabolite must be present in abundance in darkened leaves to be effective. This feature certainly limits any value this compound may have as a potential herbicide. But an understanding of the biochemistry and functional role of this inhibitor may suggest new areas for genetic manipulation of photosynthesis. □

Christine Foyer is at the Research Institute for Photosynthesis, University of Sheffield, Sheffield S10 2TN, UK.



100 years ago

I VENTURE to call attention to a simple and effective way of demonstrating the linear expansion of solids when heated, first suggested, I believe, by M. Kapoustine (*Journal de Physique*, December 1883, p. 576). The principle is, to magnify the slight extension of a bar by causing the end of it to roll upon a needle, and thus turn the latter round and move a pointer attached to it through a sensible arc.

A small flat rod of the material to be examined is laid upon two wooden blocks, placed about 25 cm apart. A weight is put upon one end of the rod to keep it from moving; under the other end, at right angles to the length of the rod, is laid a fine sewing-needle, to the eye-end of which a light pointer of straw, about 16 or 20

cm long, is attached by sealing-wax. Behind the pointer (which is painted black) a screen of white cardboard is fixed on the wooden block by drawing pins.

Before the experiment is shown to an audience, it is well to make sure that the needle rolls fairly and freely between the bar and the block. Such precautions, however, are not in the slightest degree necessary for school-work; for there is always one thing which gives the typical boy greater pleasure than to see an experiment succeed, and that is — to see it fail.

H.G. MADAN

Eton College



From *Nature* **35**, 89; 25 November 1886.

Community ecology

Patterns in three dimensions

Paul H. Harvey and John H. Lawton

MUCH of the best in ecology comes from recognizing and correctly interpreting patterns of community structure. Occasionally, the interpretation of one pattern can throw light on another, even in an apparently disparate area of the science. Just such a claim is made by Brown and Maurer on page 248 of this issue¹. They find that within a group of species that exploit the same class of environmental resources in a similar way, larger-bodied species seem to use up a disproportionate amount of the resources available. They go on to argue that, if their results are correct, a mixture of individual and species selection could provide an explanation for Cope's rule—the finding that members of taxonomic groups typically become larger sized over evolutionary time².

Brown and Maurer's analyses concern the relationship between population density and body size in assemblages of birds, fishes, rodents and plants. Larger bodied species are found at lower population densities. Damuth earlier pointed out that population density among African ungulates decreases with approximately the -0.75 power of body size³. Because mammalian basal metabolic needs increase with the 0.75 power of body size⁴, it seemed to follow that populations of large and small species were using about the same amount of energy. Peters⁵ subsequently claimed that the exponents relating population density to increases in body size over a wider range of animal groups are less than -0.75 , probably nearer -1.0 , indicating that smaller bodied species use more resources. In contrast, the latest analyses by Brown and Maurer suggest that across species exploiting similar resources in a similar way within communities, exponents are appreciably shallower than -0.75 . Hence, large species seem to use more energy per population than do small species.

Population density and body size can be considered with a third community variable, the number of species. Each of the three pairs of variables—population density and body size^{3,5}; population density and number of species^{6,8}; and body size and number of species⁹—are generally considered independently of each other. But knowledge of any two of these relationships automatically provides information about the third (J.H.L. & S.C. Pimm, in preparation). Indeed, if we understand the ecological and evolutionary determinants of any two of the patterns, then the third follows as a necessary consequence. With one exception⁹, theories that have been developed to explain each separate

relationship in this three-dimensional space have not recognized their interdependence. Even if the energy flow per species is lower in small organisms, because there are almost always more small species than large species in communities¹⁰, the proportion of total community energy flow through small species may nevertheless exceed community energy flow through large species.

The problem then is to explain why energy, or resources in general, are partitioned by many small, or a few large species, and how shares are allocated¹¹. Brown and Maurer's observations on population energy use are an important step in this direction. Ecologists need to pay much more attention to establishing reliable empirical relationships between body size, population abundance and number of species in plant and animal communities, as well as to their theoretical interrelationships.

An example comes from the study of species abundance curves that describe the relationship between population density and the number of species. For many years it has been recognized that frequency histograms of the relative abundances (population densities) of the different species exploiting similar resources, when plotted with the abundance on a logarithmic scale, are normally distributed. The distribution is said to be log normal. Furthermore, the number of species in the sample often seems to bear a particular

relationship to the variance of the distribution¹⁰. Log normal distributions with this property are termed 'canonical'. Models of the ways resources might be partitioned among the different species have been produced, the most notable being MacArthur's broken stick¹¹ and Sugihara's sequential niche division⁸ models. The latter actually predicts a species abundance distribution very similar to that observed (it is both log normal and canonical). But such models rest on the assumption that individuals of the different species have identical resource requirements. As Brown and Maurer and others have shown, this assumption does not hold and, furthermore, there is usually an inverse relationship between individual resource requirements and population density. One consequence of the distribution of relative species abundances is that energy may be partitioned among species according to a log normal distribution, but that distribution will not be canonical¹². It may be that, for once, ecologists have less to explain than they originally thought. □

1. Brown, J.H. & Maurer, B.A. *Nature* **324**, 248 (1986).
2. Stanley, S.M. *Evolution* **27**, 1 (1973).
3. Damuth, J. *Nature* **290**, 699 (1981).
4. Kleiber, M. *The Fire of Life* (Wiley, New York, 1961).
5. Peters, R.H. *The Ecological Implications of Body Size* (Cambridge University Press, 1983).
6. Preston, F.W. *Ecology* **29**, 254 (1948).
7. May, R.M. in *Ecology and Evolution of Communities* (eds Cody, M.L. & Diamond, J.M.) 81 (Harvard University Press, 1975).
8. Sugihara, G. *Am. Nat.* **116**, 770 (1980).
9. May, R.M. in *Diversity of Insect Faunas* (eds Mound, L.A. & Waloff, N.) 188 (Blackwell, Oxford, 1978).
10. Preston, F.W. *Ecology* **43**, 185; 410 (1962).
11. MacArthur, R.H. *Am. Nat.* **94**, 25 (1960).
12. Harvey, P.H. & Godfray, H.C.J. *Am. Nat.* (in the press).

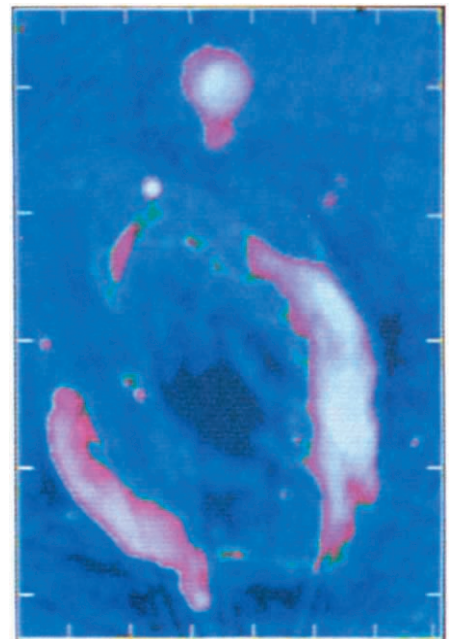
Paul H. Harvey is in the Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, and John H. Lawton is in the Department of Biology at the University of York, Heslington, York YO1 5DD, UK.

A new nebula

A COLOUR-ENHANCED image of the Circinus X-1 region of the galactic plane obtained with the Sydney University Molonglo Observatory telescope at 843 MHz, with Circinus X-1 itself in the upper region of the image.

Circinus X-1 may well be one of the more peculiar cosmic objects. It appears to be a binary system, one part of which is so compact it may be a neutron star or a black hole.

As discussed by R.F. Haynes *et al.* on page 233 of this issue, the association of Circinus X-1 with with newly discovered radio nebula is in contrast to most X-ray emitting binary stars. Circinus X-1 is in some ways similar to Cygnus X-1, but no analogous extended radio component has been discovered associated with Cygnus X-1. Comparison is also made with SS433 and the Crab nebula. It is suggested that the nebula results from accumulated energetic particles emitted during flaring activity.



Solar System

How fast does Halley spin?

Philip Campbell

ASTONISHINGLY, the rotation rate of the nucleus of comet Halley has suddenly become the focus of intense debate, eight months after the famous spacecraft encounters with the comet. The fact that significant strides have been achieved in the analysis of spacecraft and ground-based photometry and spectroscopy was clear from a recent meeting*. But above all, many participants still seemed to be recovering from the announcement at the previous week's Halley symposium in Heidelberg of the detection of a period in light curves more than three times longer than that hitherto adopted for the nuclear rotation.

The simplest way to investigate cometary rotation, given the acute difficulty in resolving cometary nuclei from Earth, is to seek regular variations in brightness. Because of the vagaries of the allocation of telescope time, there are no light curves available that are continuous since the comet's first detection, making an unbiased Fourier or maximum entropy analysis impractical. One extended collection of data was discussed by M. J. S. Belton (Kitt Peak) but, in reporting a period of 54 hours, he emphasized that he was guided by the selection of possible periods by two previous results — observations from the Japanese spacecraft Suisei of a 2.2-day (53-hour) variability in the ultraviolet intensity of the vast hydrogen cloud surrounding Halley, and an analysis by Z. Sekanina and S. Larson of 1910 dust jet images.

The more sophisticated way to measure the rotation rate of a nucleus is to send three spacecraft to look at it. On 6 March, Vega 1 saw the long axis of the nucleus pointing in a Sunward direction. Three days later, Vega 2 passed Sunward of the nucleus and obtained a broadside image. The two images were consistent with a 480 degree rotation about a spin axis perpendicular both to the long axis of the nucleus and to the orbital plane, as would accord with a 53-hour period. More recently the Giotto and Vega images were considered in combination by both imaging teams who independently arrived at a 53-hour period.

For much of the Heidelberg meeting the cometary scientists had been happy with this rotatory paradigm. But the harmony was broken somewhat when R. L. Millis and co-workers from the Lowell Observatory announced the detection of a strong 7.4-day period in ground-based photometric observations using emission bands

such as C₂, CN and OH (*Nature*, in the press). The light curves, measured in March and April this year, reveal a characteristic double-peaked structure which Millis identified as that produced by active regions entering sunlight once every rotation period.

No-one questions the reality of the 7.4-day period in the data — it has now been reported by other observers. For example M. F. A'Hearn *et al.* (Maryland), following their unexpected discovery of the existence of spiral jets of CN thought to be the photodissociation product of dust emissions (*Nature*, in the press), have found that a 7.4-day period and not a 2.2-day period satisfies the relative orientations of the jets at different times. In contrast, B. Smith (Arizona) of the Vega imaging team is adamant that the orientations of the nucleus at the three spacecraft encounters cannot accommodate such a period. It is all the more intriguing, therefore, that another member of the Vega team, J. L. Bertaux (Service d'Aéronomie, Verrières-le-Buisson) is less categorical. Nevertheless it does seem that proponents of the 7.4-day rotation period, if they are to conform to the imaging constraints, will have to appeal to a rotation axis more inclined to the long axis of the nucleus and a more complicated nuclear geometry than has hitherto been thought necessary.

As it happens there is still scope for further image enhancement. If the 2.2-day rotation period survives such scrutiny, what can be made of the 7.4-day light curve variations? Their regularity points to some dynamical origin and that most widely assumed seems to be a free wobble (a motion not maintained by a torque) equivalent to the terrestrial Chandler wobble of the Earth, though much closer to the rotation rate of the comet than is the rate of terrestrial Chandler wobble to the rotation rate of the Earth. At Heidelberg, K. Wilhelm of the Max Planck Institute, Lindau, who had analysed the motion of an appropriate-sized ellipsoid, reported that a wobble period of about six days was not unreasonable and was consistent with the spacecraft images.

Despite the fact that such a wobble could not be observed directly from outside the cometary coma, its behaviour, if manifested in photometric light curves, could yield useful information about the interior of the nucleus. The wobble would presumably be excited by the jet activity as the comet passed through the vicinity of perihelion. Over the subsequent orbital period, as shown by Wilhelm, it should be damped out by internal friction. But in

Paris, M. Festou (Observatoire de Besançon) presented an amalgamation of all available photometric data taken between January 1984 and February 1985, showing that a 7.4-day period was significant throughout. As J. Lissauer (U. C. Santa Barbara) emphasized, interpreting this as a wobble would imply a significant amplitude near the very onset of emissions from the nuclei, in turn implying small internal damping and a comparatively rigid nuclear interior.

Although such detailed studies of the behaviour of Halley have proved unexpectedly engrossing — and may remain so for some time — more wide-ranging questions were also addressed in Paris, relating in particular to the vast but putative Oort cloud of comets from which Halley is thought to have originated. The darkness of cometary nuclei, emphasized by spacecraft images of Halley, combined with the new findings of the shape and degree of activity of the nucleus, allowed P. Weissman (Jet Propulsion Laboratory) to generalize and estimate afresh the total mass of the cloud, increasing it over a thousandfold above Oort's original estimate of a tenth of an Earth mass or less. This could imply that more comets have collided with Earth than hitherto thought, also strengthening the idea that an early generation of comet-like objects may have delivered the stuff from which the atmospheres and oceans to be found on the terrestrial planets are derived.

The perturbations of the Oort cloud thought to produce long-period comets have traditionally been ascribed to random encounters with passing stars. Recent theoretical work, however, indicates that ever-present galactic tidal forces should also play a significant role. A.H. Delseume and M. Patmiou (Toledo) plotting the aphelia of 152 selected long-period comets in galactic coordinates, reported 'zones of avoidance' at the Galactic poles and equator, consistent with a tidal effect.

In reviewing the progress in the analysis of spacecraft data, E. Grün (Max Planck Institute, Heidelberg) described surprising results from the PUMA dust impact mass analysers on the Vega spacecraft. Small particles (< 1 µm) have been found with relative abundances of C, N and O very similar to those in the Sun and much higher than those found in carbonaceous chondritic meteorites. Similarly, gaseous isotope ratios measured from spacecraft and from the ground are consistent with solar abundances. Though the spacecraft results are still preliminary, all the signs are that — as has always been hoped — comets contain by far the most primitive and pristine material to be found in the Solar System. □

Philip Campbell is Physical Sciences Editor of *Nature*.

*Annual meeting of the Division of Planetary Sciences of the American Astronomical Society, Paris, 4–7 November 1986.

Yeast genetics

Genetic control mechanisms: transcriptional twisting

Ian W. Dawes

YEAST geneticists and molecular biologists have long been advocating yeast as a model eukaryotic organism as well as an organism for exploitation by the biotechnologist. Reservations about the usefulness of such a model are based largely on the relatively uncomplicated nature of the yeast life cycle and partly on differences from higher eukaryotes in the mitotic and meiotic processes. At a recent meeting*, however, it became clear that much progress has been made in the study of those cellular processes for which yeast is ideally suited. One example of such a process is the regulation of transcription.

Now that important regulatory sequences have been identified upstream of several genes in yeast, the search is on in earnest to find proteins binding to these sequences and to determine how this binding leads to specific control as well as initiation of transcription. With regard to the nature of the upstream sequences concerned with positive control, the seven genes for allantoin degradation (*DAL1* to *DAL5*, *DAL7*, *DUR1,2*) which are all under the control of the *DAL81* gene, contain two types of consensus region (T.G. Cooper, University of Tennessee, Memphis). For efficient constitutive expression, Cooper proposed an upstream expression sequence (UES) which is found in front of all the above genes at least once. For the highly regulated *DAL4*, *DAL7* and *DUR1,2* genes there is a second sequence, an upstream modulation sequence (UMS) present in multiple copies, which is involved directly in the inducibility of these genes. For the less inducible genes (*DAL1* and *DAL2*) there are fewer UMS elements, and these exhibit only limited homology with the consensus region. For the constitutive genes (*DAL3* and *DAL5*) there is no UMS.

The distinction between expression and regulation sequences is not radically different from what has been reported for other genes that have been examined (and maybe soon the terminology will be agreed), but the nice point of this work is the extent to which the correlation between control sequences and inducibility holds for a group of genes under one type of control.

It also illustrates the extent to which UMS and UES sites can be fairly scattered around in the 5' upstream region of a naturally regulated system. For negative control of the *CAR1* gene, Cooper's group

has found a 13-base pair sequence (an upstream repression sequence, URS) that is necessary for repression, and which can control *CYC1* expression if cloned upstream or downstream of the upstream activating sequence (UAS) elements. This sequence is present in the 5' flanking regions of more than a dozen genes of widely diverse functions, including nitrogen catabolism, protein synthesis and glycolysis, and it may be part of a more global regulatory system, or possibly a generic sequence associated with a repressor binding site as it does show some homology with the phage λ repressor binding site.

Several points have emerged about proteins binding at the UAS elements. The DNA-binding proteins that have been sequenced have short domains for DNA binding relative to the overall length of the molecules (for example, the amino-terminal 74 residues from 881 for the product of the regulatory *GAL4* gene: (G. Gill, Biochemistry and Molecular Biology, Harvard University) and 60 carboxy-terminal residues from 282 for the *GCN4* gene product conferring general amino-acid regulation (K. Struhl, Harvard Medical School). These DNA-binding regions show the helix-turn-helix motif of prokaryotic repressors. Other regions of these control proteins are necessary for the activation of transcription, thus fusions of the carboxy-terminal end of the *GAL4* gene product and the *Escherichia coli* *lexA* protein will bind to *lexA* promoters and initiate transcription in a heterologous yeast system (Gill). The elegant work from Struhl's laboratory with the *GCN4* gene product shows that in addition to the short DNA binding site at the carboxy terminus, this protein has a central and even shorter (19-residue) acidic domain needed for the activation function. There are similar acidic regions in other yeast regulatory proteins and in the HMG proteins thought to be important in regulation of higher cell development (Struhl). Deletions between the DNA-binding and acidic domains do not abolish the activity of the protein *in vivo*.

The presence of a single DNA-binding regulatory protein at the UAS is not sufficient for control — other proteins are involved. This has been found for the upstream control of the two cytochrome *c* genes *CYC1* and *CYC7* (L. Guarente, Massachusetts Institute of Technology). *CYC1* expression is regulated by haem and glucose from two elements, *UAS1* and *UAS2*. Dissection of the 73-base pair

UAS1 region reveals one protein (plus) binding the one region, and two others (*HAP1* and *RC2*) to another. How do the proteins binding at these UAS sites affect the initiation of transcription which can be located from 250 to 1,000 or more base pairs downstream from the UAS site? Rudi Planta (Vrije Universiteit, Amsterdam) described the proponents of various hypotheses as being either "twisters, sliders, oozers or loopers" (see Ptashne, M. *Nature* **322**, 697–701; 1986). Gill, a confirmed looper, suggested for the *GAL1* gene that up to four molecules of the *GAL4* product bind at *UAS_G* and interact with other proteins binding nearer the start of transcription, thereby introducing a loop in the DNA and bringing the UAS and the transcription start site into juxtaposition.

With all the current discussion of whether the human genome should be completely mapped and sequenced, and who will pay for the exercise, it is interesting to see how far similar exercises have progressed in an organism with only a fraction of the haploid DNA content. M. Olson (Washington University School of Medicine, St Louis) reported on progress in the task of restriction-mapping the entire yeast genome at the 2-kilobase level. By using 'bottom up' screening of about 5,000 λ inserts for overlaps, about 600 local maps have been identified, which should reduce to about 200 on a second pass through the restriction data. The other approach of 'top down' mapping using OFAGE electrophoresis to separate the 16 yeast chromosomes and map them individually with the rare cutting restriction enzymes *NotI* and *SfiI*, generating about 200-kilobase fragments, is also well under way, with more than half of the 55 *SfiI* fragments mapped, and complete maps for eight of the chromosomes. One of the early goals of this work should soon be realized: the production of an ordered set of λ clones to allow precise location of a newly cloned gene in a single hybridization experiment. The organization of chromosome 1 (the smallest chromosome, in which 250 kilobases can encode only about 100 genes) is the furthest advanced. Kaback and colleagues (New Jersey Medical School, Newark) have mapped 160 kilobases of almost contiguous DNA representing 70 per cent of the chromosome, and have also studied transcription from genes previously undetected by mutation.

Any report of a meeting of the size and interest of this one is bound to be fragmentary and selective. Fortunately, for the first time the complete abstracts of a meeting in this series are available in published form as a supplement to the journal *Yeast*. □

Ian W. Dawes is in the Department of Microbiology, University of Edinburgh, West Mains Road, Edinburgh EH9 3JG, UK.

*13th conference on yeast genetics and molecular biology, held at Banff, Alberta, 1–5 September 1986.

Fingers and helices

SIR—On several occasions the interaction between *Xenopus* transcription factor IIIA (TFIIIA) and 5S RNA genes has been discussed in *Nature*, *News* and *Views*¹⁻³. TFIIIA is a good example of the so-called finger proteins⁴, eukaryotic regulatory proteins which contain variable numbers of a characteristic tandemly repeating motif of about 30 amino acids^{5,6}. It is currently thought that these sequence repeats represent small nucleic acid binding domains, each folded about a tetrahedrally coordinated zinc ion. The coordination is provided by appropriate side chain atoms from fully conserved histidines and cysteines. This had led to a model in which TFIIIA is essentially a 'string of beads' that zigzags along one side of a nucleic acid double helix in an extended fashion⁷.

Our observation⁵ of the repeating motif in TFIIIA, which has unfortunately been overlooked by most authors, allowed us to make a reasonable prediction of the average secondary structure of the repeat. We have now extended the secondary structure analysis to include all the finger sequences known to us (Table 1). The prediction technique⁸ involves calculation of smoothed plots for each of the three conformational preference parameters (α -helix, β -strand and turn) for individual sequence repeats. The smoothed curves are then averaged over all the repeats. The most strongly predicted region spans alignment positions 17–23 (Table 1) where the mean smoothed helix preference varies from 1.1 to 1.15. Given that, 1.0 indicates a neutral preference and that 1.4 is the largest preference, the helical

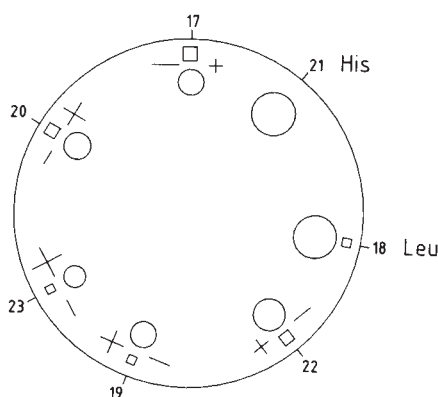


Fig. 1 A 'helical' wheel representation of the predicted α -helical region in the 'finger' sequences. The wheel is numbered according to the alignment in Table 1. The positions of the strongly conserved Leu 18 and His 21 are shown. The symbols used to illustrate the amphipathic nature of the α -helix are (○) for hydrophobic residues (IALVMCHYWF), (+) for basic amino acids (KR), (-) for acidic/polar residues (QEDN), and (□) for (PSTG). The size of the symbols are proportional to the amino acid composition at each alignment position.

Table 1 Aligned 'finger' sequences

*Repeat	1	10	20	30	*Repeat	1	10	20	30
1	YICSFADCGAAYNKNWKLQAHLC-KHTGEKP-				24	HICPI--CGVIRRDDEEYELHLMN-LHEGKTE-			
2	FPCKEEGCEKGFSTLHHLTRHSL-THTGEKN-				25	KQCRY--CPKFSRFPVNTLRHMR-SHWDKKK-			
3	FTCDSDGCDLRFTHKANMKHFNRFHNIKCV				26	YQC--EKCGLRFSQDNLLYNHRL-RHEAENP			
4	YVCHFENCGAKFKHNLQKVHQP-SHTQQLP-				27	IICSI--CNVSFKSRKTFNHTL-IHKNRFR			
5	YECPHGEGDKRFSPLSRKLKHEK-VH--AGY				28	HYCSV--CPKSFTERYTLKMHMK-THGDDVY			
6	PCKKDDSCSFVGKWTLYLKHVAECH-QDL--				29	F-CLI--CNTTFENKKELEHHLQFDH--DVS-			
7	AVC--DVCNRKFRHKDYLRDHQK-THEKERTV				30	LHCR--CRTQFSRRSLHLHQLKRCQDQFSV			
8	YLCPRDGCDSRYTAFNLRSHIQSFHEEQRP-				31	FVC--EVCTRAFAEQELKRHYR-SHTNEKP			
9	FVCEHAGCGCFAMKKSLEHRSV-VHDPEKPK				32	YPCGL--CNRCFTFRDLIRHAQKIHSGNLGE			
10	FTCKI--CSRSGFYKHLQNHQR-THTGEKP-				33	FVC--NYCDKTFSEKSLVSHKR-IHTGEKP-			
11	FECPE--CDKRFTRDHLKTHMR-LHTGEKP-				34	YEC--DVCQKTFSEKSLVSHKR-IHTGEKP-			
12	YHCSH--CDRQFVQVANLRHRL-VHTGERP-				35	FECPE--CGKFTAQKSLVSHKR-AHMEKFP-			
13	YTC--EICDGKFSDSNQLKSHML-VHTGEKP-				36	YGCSE--CGKFTAQKSLVSHKR-IHTGEKP-			
14	FEC--ERCHMKFRRRHHLMNHKCGIQSPPTPA				37	YEC--NECAFTFFKSLVSHKR-IHTGEKP-			
15	FEC--EFCHKLFVSKEDLQVHRR-IHTKERP-				38	YECSE--CGKFSIQNSQLIHRH-IHTGEKP-			
16	YKC--DVCGRFAFESGLRHRMR-IHTGERP-				39	YECTE--CGKFSIQNSQLIHRH-IHTGEKP-			
17	IPCHI--CGEMFSSQEVLEHRIKADTCQKSEQ								
18	ATC--NVCGGLKVKDDEVDLRHME-LHEGKTE-								
19	LECRY--CDKFSHKKRNVLRHME-VHWDKKK-								
20	YQC--DKCGERFSLSWLMYHNLH-RHDAEENA								
21	LIC--EVCHQGFQKTKRTYKHLRL-THQDTRPR								
22	YPCPD--CEKSFVDKYTLKVKHR-VHQPVEKP								
23	QECTT--CGKVYNSWYQLQKHISEHESKQ-PN								

*Repeats 1–9 are from *Xenopus* TFIIIA. 10–14 are from *Drosophila Kt*, 15–16 from *Krh* (R. Schuh and H. Käckle), 17–22 from beta *sry*, 23–29 from delta *sry* and 30 from IB142 (R. Baldarelli and J. Lengvel). 31–32 are from yeast ADRI and 33–39 from mouse MKR3 (K. Chowdhury and P. Gruss).

†The consensus for an alignment position is given if 19 or more of the 39 residues are conserved according to the groups: (ST), (PG), (KR), (QEDN), (MC), (HFYW) and (AIVL).

prediction over the 39 repeats is clear.

An α -helical wheel representation⁹ of the span (Fig. 1) shows an obvious hydrophobic/hydrophilic sidedness, typical of helices observed in many three-dimensional protein structures. One of the conserved histidines suggested for zinc coordination and a strongly conserved leucine are located on the hydrophobic side of the predicted α -helix. This span may act as the principal finger region for interaction with nucleic acid.

Two preliminary spectroscopic observations in this laboratory support the helical suggestion. Circular dichroism curves of the TFIIIA protein as analysed by the method of Provencher¹⁰ indicate the presence of 10–15 per cent α -helix. A fluorescence quenching analysis of the 7S particle shows an approximate 30 per cent increase in intensity upon digestion of the RNA with ribonucleases or its salt-induced dissociation. We interpret this re-

sult as evidence for changes in the environments of the two tryptophan residues (W28 and W177) in TFIIIA which are just at the amino-terminal side of the predicted α -helix. Apparently their degree of exposure and structural flexibility is altered considerably upon binding to nucleic acid.

RAYMOND S. BROWN

PATRICK ARGOS

European Molecular Biology Laboratory,
Meyerhofstrasse 1,
D-6900 Heidelberg, FRG

1. Enver, T. *Nature* **317**, 385–386 (1985).
2. Berg, J.M. *Nature* **319**, 264–265 (1986).
3. Harrison, S.C. *Nature* **322**, 597–598 (1986).
4. Vincent, A. *Nucleic Acid Res.* **14**, 4385–4391 (1986).
5. Brown, R.S., Sander, C. & Argos, P. *FEBS LETT.* **186**, 271–274 (1985).
6. Miller, J., McLachlan, A.D. & Klug, A. *EMBO J.* **4**, 1609–1614 (1985).
7. Rhodes, D. & Klug, A. *Cell* **46**, 123–132 (1986).
8. Argos, P. *EMBO J.* **4**, 1351–1355 (1985).
9. Schiffer, M. & Edmundson, A.B. *Biophys. J.* **7**, 121–135 (1967).
10. Provencher, S.W. & Glöckner, J. *Biochemistry* **20**, 33–37 (1981).

Data in graphs and tables

SIR—After reading Paolini's discussion on enzyme units and how to express them¹, my attention was drawn to a similar ambiguity in the literature, concerning the units in which data are expressed when shown in tables or plotted in graphical form.

Basic mathematics tells us that the value of a physical quantity is equal to the product of a numerical value and a unit:

physical quantity = numerical value $\times$ unit

According to the convention followed by the Royal Society², the expression used to define the numerical values of a physical quantity plotted on a graph (or the one which is placed at the head of a column of numerical values of a physical quantity in

a table) should be a pure number, and the axis on the graph (or the head of the table) should be labelled with physical quantity/unit — for example, radioactivity/(d.p.m. $\times 10^3$) — thus, making the numbers in graphs and tables dimensionless. If this convention is not followed, that is, if the numerical values in graphs and tables are considered to be dimensional, then the axis on graphs and the heads of tables should be labelled with the appropriate units (d.p.m. $\times 10^3$, and not d.p.m. $\times 10^{-3}$ as is often the case).

Figure 1 summarizes several ways of expressing data in graphical form. In this figure, molar was the chosen unit because conventional prefixes like nano, micro, milli, and so on, are very commonly used, which is not the case with d.p.m. (for which Bq is now the recommended unit)

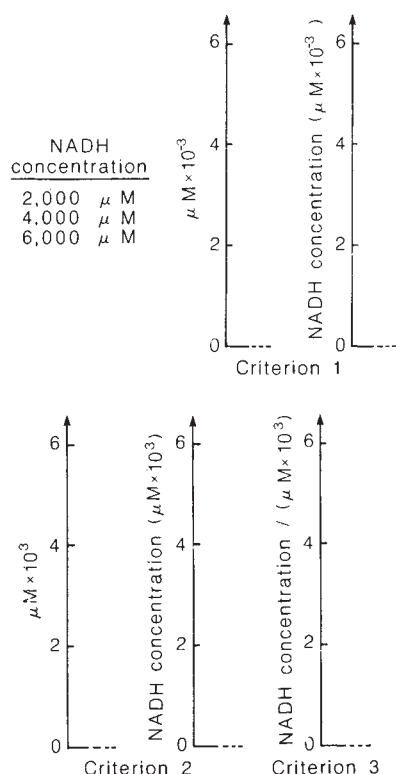


Fig. 1 Hypothetical experimental data showing common criteria for graphical presentation.

and some other units. Despite the method approved by the Royal Society (criterion 3) being mathematically the most precise and elegant, it is difficult to envisage its general use particularly outside the scientific community (in, for example, newspapers and television). In many (most?) biochemical papers, criterion 1 seems to be followed and even in well established books it is easy to find inconsistent usage: here for example, in ref. 3 criterion 1 is used in fig. IV-29 but fig. IX-2 uses criterion 2. Use of criterion 1 may be prone to ambiguous interpretation as shown in Fig. 1 — with data in thousands of μM (indeed mM), we end up with the axis labelled $\mu\text{M} \times 10^{-3}$ when this criterion is followed (not in fact mM , but nM !).

Although in virtually all cases it is easy to recognise which method has been used (and this may explain why several methods coexist), the indiscriminate use of different ways of labelling graphs and tables does not contribute to the clarity with which scientific knowledge must be presented. In this respect, I would rule out criterion 1 because it is the most ambiguous.

RICARDO FERREIRA

Department of Biochemistry,
School of Biological Sciences,
University of East Anglia,
Norwich NR4 7TJ, UK

1. Paolini, M. *Nature* **321**, 568 (1986).
2. The Symbols Committee of the Royal Society *Quantities, Units and Symbols* (The Royal Society, London, 1971).
3. Dixon, M. & Webb, E.C. *Enzymes* (Longman, London, 1979).

Silver from Chernobyl

SIR—I beg to differ with R.H. Flowers (*Nature* **323** 208; 1986) concerning the origin of $^{110\text{m}}\text{Ag}$ in fallout from Chernobyl.

It is easily calculated that at most 400 grams of the fission product ^{109}Ag had been produced in the reactor core of Chernobyl-4 up to the moment of the accident. On the other hand the reactor was equipped with about 200 in-core neutron detectors of the self-powered type with emitters composed of silver. These emitters are up to 2.5 metres per detector and, in all, contain many kilograms of natural silver, about 48 per cent of which is ^{109}Ag .

Therefore the $^{110\text{m}}\text{Ag}$ which has also been detected in significant amounts in the Netherlands, is indeed an activation product but almost entirely originates from molten detectors.

H. VAN DAM

Interuniversity Reactor Institute,
2629 JB Delft, The Netherlands

Chernobyl fallout in Debrecen, Hungary

SIR—Since 1952, total beta activity of fallout in precipitation has been measured in Debrecen without interruption^{1,2}. The precipitation is collected daily by a PVC funnel (40 cm diameter) and evaporated into standardized glass cups, where total beta activity is measured by an end window Geiger-Müller counter.

On 29 April, 2.7 mm (340 cm^3) of rain was collected, but previous activity levels were unchanged. However, the shower of

rain on 1 May (0.48 mm) resulted in 1,450 Bq m^{-2} initial total beta activity (Fig. 1). The next rainfall, on 9 May, was collected and processed in two parts (2.5 mm before noon and 3.1 mm afterwards). The initial activity of the first part was 2,500 Bq m^{-2} and that of the second part was only 160 Bq m^{-2} , which is impressive proof of the rapidity of change in tropospheric radioactivity and shows that the main mechanism of fallout deposition is rain-fall. The specific activities of these three samples were 2,645; 980 and 46 Bq l^{-1} respectively.

We have identified the radioisotopes by the γ -ray spectra of the samples using a $100 \text{ cm}^3 \text{ Ge (Li)}$ detector surrounded by a 5 cm lead shield. The isotopic composition of the fallout in the precipitations before noon on 1 May and 9 May proved to be the same, though the intensity ratios of short half-life radioisotopes were different.

The residual activity on the later date was that of the isotopes ^{103}Ru , ^{95}Zr , ^{95}Nb , ^{106}Ru , ^{106}Rh and $^{134,137}\text{Cs}$.

In all, after 10 May the total beta activity level was less than in the first three samples, and did not significantly exceed, those of past years after 25 May.

E. CSONGOR

A.Z. KISS

B.M. NYAKÓ

E. SOMORJAI

Institute of Nuclear Research,
Hungarian Academy of Sciences,
H-4001 Debrecen,
POB 51, Hungary

1. Szalay, A. & Csongor, E. *Acta phys. hung.* **29**, 407–413 (1970).
2. Csongor, E. *Fizikai Szemle* **39**, 11–20 (1985).

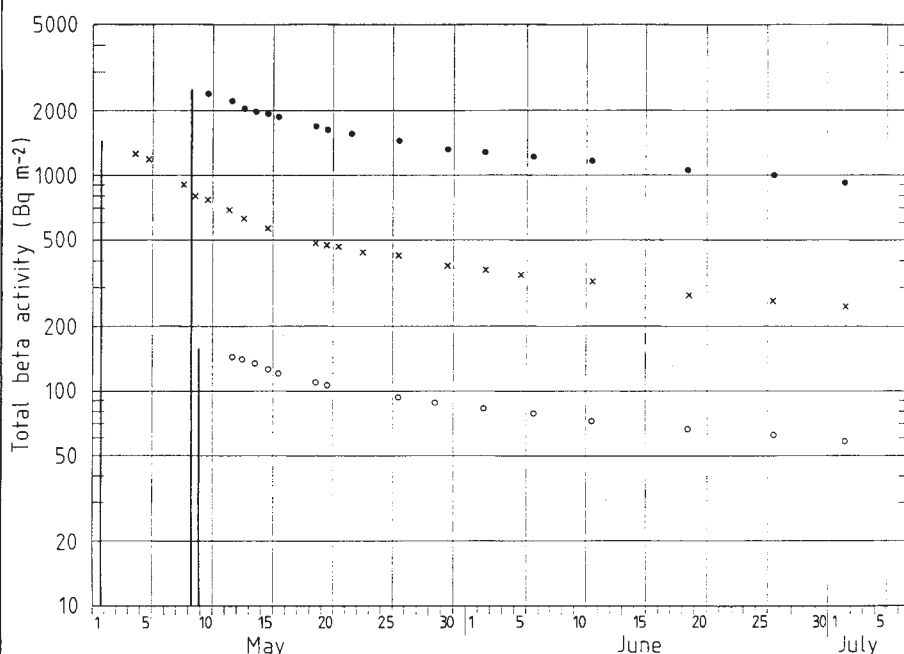


Fig. 1 Total beta activity of the Chernobyl fallout in the precipitation in Debrecen, Hungary during May–July 1986.

Birth of a separate science

Peter Harman

Intellectual Mastery of Nature: Theoretical Physics from Ohm to Einstein. Vol. 1 The Torch of Mathematics 1800–1870; Vol. 2 The Now Mighty Theoretical Physics 1870–1925.

By Christa Jungnickel and Russell McCormmach. *University of Chicago Press: 1986. Vol. 1 pp. 350, \$55, £46.75; Vol. 2 pp. 434, \$65, £55.25.*

IN AN autobiographical sketch, written in 1891 towards the end of his exemplary career, Hermann von Helmholtz explained that he had sought “intellectual mastery of nature” in pursuing a career in science. Originally working principally in physiology, in physiological optics and acoustics, he had also made substantial contributions to physics. Helmholtz’s classic memoir *On the Conservation of Force* (1847) provided a mathematical formulation of the principle of the conservation of energy, one of the grand unifying principles of nineteenth-century physics; and papers on hydrodynamics and electrodynamics confirmed his reputation as one of the greatest mathematical physicists of the day. Relinquishing physiology, he moved to Berlin in 1870 as professor of physics; he found the subject to be “the true basis of all proper science”.

It was in theoretical physics, above all, that German scientists found “science as a vocation”, and it was through the “torch of mathematics” (in Ohm’s telling phrase) that they sought to realize their intellectual mastery of nature. At the beginning of the nineteenth century, German universities provided no more than elementary physics lectures for students who would go on to professional studies in medicine, law or theology. The main argument for educational reform was the hoped-for regeneration of the cultural and political reputation of the German states; its tool was the incorporation of research into the universities. The effect was the creation of scientific institutes and increasing specialization within the natural sciences. In their superbly written and exhaustively researched study, Jungnickel and McCormmach find the division of physics into theoretical and experimental specialities to be the distinctive feature of German physics in the nineteenth century.

Drawing on a rich supply of archival materials from universities and state bureaucracies, together with a thorough sampling of the physicists’ scientific work, Jungnickel and McCormmach present a comprehensive and convincing picture of physics in nineteenth-century Germany. Some very complex science is described with economy and precision, but without undue technicality, and the authors have written an integrated narrative in which

the scientific work of the physicists is portrayed in its institutional setting. The collaboration has proved remarkably successful, with the authors’ interests (Jungnickel’s more in institutional history, McCormmach’s more in intellectual

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Hermann von Helmholtz, “classical master”.

history) being effectively blended. In his *Night Thoughts of a Classical Physicist* (Harvard University Press, 1982) McCormmach has already presented a glimpse of the materials here meticulously documented, in the form of a fictional reconstruction of the life of a German physicist active from the latter part of the nineteenth century.

Physical theory, or “natural philosophy”, had long been a subject of study, yet only by the beginning of the nineteenth century was the term “physics” being used in the modern sense to denote a study of mechanics, optics and electricity which employed mathematical and experimental methodology. At that time physicists strongly emphasized quantification and the search for mathematical laws, and sought a unified science of physics based on the concept of energy and the programme of mechanical explanation. In seeking to give an integrated account of the scientific work in its institutional setting, Jungnickel and McCormmach concentrate on the physicists’ work as an organizing principle. Research and teach-

ing were carried out in institutes within the universities, and the book is structured around the development and proliferation of these institutes.

The pattern of development was remarkable. Early in the nineteenth century German physics was struggling to distance itself from the prevalent “nature philosophy”. University professors and authors of texts tried to combat Hegel’s contempt for the natural sciences. Around 1830 physics began to become established as an academic discipline, especially at Göttingen and the new university of Berlin. Gauss’s collaboration with Wilhelm Weber in Earth-magnetic observations, Ohm’s mathematical theory of electric currents and Franz Neumann’s work on optics and electricity all contributed to establishing a coherent corpus of German physics. Much of the inspiration for mathematical physics came from France, the work of Laplace, Poisson, Fresnel and Fourier providing notable exemplars, but by the 1840s a German tradition was emerging. Helmholtz’s unification of physics in terms of the conservation of energy drew on work by Gauss, Weber, Ohm and Neumann. German work on electricity was especially notable; in the 1840s, Weber, Kirchhoff and Neumann published substantial memoirs. As a tyro electrician in 1854, James Clerk Maxwell felt the need to study these works, the “views of heavy German writers”. Weber’s electrodynamics, Helmholtz’s energy physics, Clausius’s work on thermodynamics and the “mechanical” theory of heat, all strove for unifying concepts by which the foundations of physics could be established.

The subject prospered even though physics institutes suffered from shortage of space and resources, and after 1870 the development of theoretical physics as a study distinct from experimental physics owed much to the new teaching needs which faced physics institutes. The requirement for a second teaching post within individual universities led to the creation of “extraordinary” professors in theoretical physics. These positions were planned as transitional appointments for young scientists who were expected to move on to an “ordinary” professorship in experimental physics and to the control of a physics institute. But eventually, following the example of Berlin, institutes of theoretical physics came to be established with theoretical physics being understood as that part of the subject concerned with the precise measurement of physical constants as well as with mathematical physics. Specialization was considered acceptable if the twin sides of the subject were kept in balance: as Helmholtz urged, without mathematical physics, experimental physics is a very narrowly bounded science and affords little insight into the course of physical

phenomena, while . . . mathematical physics without experimental physics, would likewise be a rather lame and unfruitful science.

This attitude shaped Helmholtz's evaluations of candidates for university positions. Jungnickel and McCormmach consider Max Planck as the first true specialist in theoretical physics (in the sense that his career developed entirely in the subject). Yet Planck was considered by Helmholtz to be an "excellent acquisition" in replacing Kirchhoff as Berlin's theoretical physicist in 1889; Planck was a "splendid man" because his work was closely based on the experiments of others, on the interpretation and development of experimental measurements.

The main intellectual influence shaping German physics after 1870 was the work of James Clerk Maxwell. Helmholtz's influential treatment of electrodynamics and evaluation of the unique importance of Maxwell's electromagnetic theory of light encouraged the research of Heinrich Hertz, first under Helmholtz's direction at Berlin and then at Karlsruhe, which ultimately led to his decisive experimental test in favour of Maxwell's theory in 1888. As a visitor to Helmholtz's institute anxious to gain experience of experimental work, Ludwig Boltzmann undertook experiments (on dielectric constants) which bore on the confirmation of Maxwell's theory. In his research on the kinetic theory of gases and the entropy concept, Boltzmann extended Maxwell's law of the distribution of velocities among the molecules of gas, creating statistical thermodynamics. And in Leiden, Lorentz developed the "electron" theory which blended elements from German electrodynamics with Maxwell's theory of the electromagnetic field. The theoretical initiatives of Lorentz and Boltzmann, both much sought-after in German institutes of theoretical physics (Lorentz remaining in Leiden, but Boltzmann seeking to assuage his depression by moving to and fro between Graz, Vienna, Leipzig and Munich), encouraged German physicists to look to Maxwell's electromagnetic theory and molecular mechanics for their inspiration.

By the end of the century German physics, especially theoretical physics, manifested classical coherence and also innovatory thrust. The lectures of the classical masters Helmholtz and Kirchhoff summarized and codified the achievements of the past. The endeavours of Planck and Wien, debating the rationale of energetic and electromagnetic foundations for physics, opened the door to the new physics of quanta and relativity of the twentieth century, itself a creation of German theoretical physicists yet shaped by the problems which were perceived to flow from Maxwell's work. Planck came to exercise a formative influence, succeeding Kirchhoff in Berlin, and on Helmholtz's death becoming the advisor for

theoretical physics for the *Annalen der Physik*. His research on blackbody radiation led him to introduce the quantum of energy, although he remained cautious about its implications. Nevertheless he was an early and influential supporter of Einstein's special theory of relativity.

Jungnickel and McCormmach conclude their account with an overview of the developments which arguably constitute the high point of German physics, the creation of the general relativity and quantum theories in the first quarter of the twentieth century. The authors emphasize the confluence of intellectual and institutional developments. Einstein came to the Prussian Academy and to Berlin in 1913 as the outstanding theoretical specialist, and

soon delivered the "scientific eggs" in his general theory of relativity. The emerging generation of theoretical physicists, including Pauli and Heisenberg, received their education at the hands of men who had pursued careers as specialist mathematical physicists, Sommerfeld at Munich and Born at Göttingen. If classical physics had achieved its apotheosis in the lectures of Helmholtz and Kirchhoff, the new generation could build creatively on the shoulders of giants. The "now mighty theoretical physics", as Wilhelm Wien described it in 1915, had attained status as a "separate science". □

Peter Harman is currently an Associate of Clare Hall, Cambridge, CB3 9AL, UK. He is author of Energy, Force, and Matter: The Conceptual Development of Nineteenth-century Physics (Cambridge University Press, 1982).

Naked appreciation

P.W. Hawkes

Principles of Charged Particle Acceleration. By Stanley Humphries Jr. Wiley: 1986. Pp.573. \$65.95, £63.15.

THE difference between electron optics and the optics of charged particle acceleration is not unlike that between nudity and nakedness. It is possible to write a book on electron optics that sets out from a few premises and provides an elegant abstract description of the behaviour of electrons in the lenses, accelerating structures and analysers of which many instruments are composed. The typical example is P.A. Sturrock's *Static and Dynamic Electron Optics* (Cambridge University Press, 1955), in which it is not too fanciful to recognize the austerity of the classical nude.

In the optics of charged particle acceleration, on the other hand, the blemishes and homely imperfections of mere nakedness cannot be overlooked: many different kinds of particle may be present in the beam, the current density may be so high that the particles jostle one another in a way that renders the fundamental equations unmanageable, instabilities may set in and, in the great accelerators of recent years, even the irregularities of the underlying terrain cannot be neglected.

Stanley Humphries is well aware of the complexity of his subject matter and steers a middle course that accelerator physicists and engineers will be able to follow easily. Like Sir Thomas Browne, he believes that "where there is an obscurity too deep for our reason, 'tis good to sit down with a description, periphrasis or adumbration". He opens with concise accounts of particle dynamics in electric and magnetic fields, field calculation and materials properties. The next three chapters are concerned

with lenses: round lenses, quadrupoles and sectors. Their effects on particle beams are first examined in terms of the elementary lens properties, after which the notions of acceptance and matching are introduced. The last of these chapters explains the transfer matrix formalism. Humphries does not go far into these optical matters, but this merely emphasizes that he has written a book on acceleration and has included only the necessary minimum about them.

Apart from a single chapter on the difficult theme of phase dynamics, the rest of the book offers a detailed and knowledgeable account of the various types of accelerator. Electrostatic accelerators and pulsed high-voltage, linear-induction accelerators, betatrons, radiofrequency accelerators and, finally, cyclotrons and synchrotrons: each of these is accorded a separate chapter, as are the resonant cavities and waveguides employed. Not only are the modes of action of the various machines clearly described, but the book is also a rich hoard of down-to-earth engineering detail that will make it used and appreciated far beyond the introductory course on the physics of charged particle acceleration that inspired it.

Humphries has brought together and successfully organized a mass of material in the first large single text to be devoted to his subject since *Linear Accelerators*, a multi-author work edited by P.M. Laposolle and A. Septier and published by North-Holland in 1970. I suspect that many readers will need to go elsewhere to delve into the subjects of the earlier chapters in more detail, but they will surely forgive the author for preferring to save his space for the accelerators themselves. Altogether, this is an impressive and consummately useful volume. □

P.W. Hawkes is Directeur de Recherche at the Laboratoire d'Optique Electronique du CNRS, 29 rue Jeanne Marvig, 31055 Toulouse Cedex, France.

How mistakes are made

Leslie E. Orgel

Accuracy in Molecular Processes: Its Control and Relevance to Living Systems. Edited by T. B. L. Kirkwood, R. F. Rosenberger and D. J. Galas. *Chapman & Hall*: 1986. Pp. 398. £48.50, \$110.

THE frequency of errors in the transmission of DNA sequence information from one generation to the next can be as low as one per 10^8 – 10^{10} bases copied. Transcription and translation have substantially higher error rates, but are still far more chemospecific than comparable non-enzymatic processes. The measurement of error rates, the study of the enzymatic mechanisms that maintain them at low levels and the investigation of the cumulative effect of errors on the health of the organisms that harbour them have become the subject matter of a subspeciality of biochemistry. This subspeciality is the subject of the present multi-author volume. The "accuracy" of methylation, restriction, splicing and other post-synthetic modifications of proteins and nucleic acids is not discussed.

The empirical data on the accuracy of DNA replication, and the classic mechanism of error correction by exonuclease activity during replication are discussed in a chapter by Goodman and Branscomb. The role of DNA repair in maintaining the integrity of the genetic material is covered by Sedgwick. The accuracy of RNA synthesis is discussed by Anderson and Menninger.

Protein synthesis is dealt with in three chapters. Fersht covers the charging of tRNA and deals in detail with editing pathways that remove "incorrect" amino acids hydrolytically from charged tRNAs. Buckingham and Grosjean review the literature on the accuracy of mRNA–tRNA recognition and show that recognition depends on many factors beyond the complementary pairing between codon and anticodon triplets. Kurland and Galant deal with the role of the ribosome in controlling the accuracy of protein synthesis.

In addition to these empirically based reviews, there are other, mainly theoretical contributions on such topics as the kinetic cost of accuracy (Ehrenberg, Kurland and Blomberg), selection for optimal accuracy (Kirkwood and Holliday) and the effects of errors on the integrity of genetic information transfer (Rosenberger and Kirkwood). A chapter on probabilistic thinking about enzyme kinetics by J. Ninio provides a number of qualitative insights that are not readily derived from the differential equations

of traditional chemical kinetics.

This volume is a valuable addition to the literature of biochemistry because it brings together reviews of all of the important aspects of the "accuracy" problem. It is not always easy going — the introductory material is highly compressed and requires very careful reading, as do many of the more theoretical sections. But the determined reader will be rewarded with an excellent overview of an important but little-known area of biology. □

Leslie E. Orgel is Director of the Chemical Evolution Laboratory, The Salk Institute for Biological Studies, San Diego, California 92138, USA.

Solar systems

A. K. Turner

Physics, Technology and Use of Photovoltaics. By R.J. Van Overstraeten and R.P. Mertens. *Adam Hilger*: 1986. Pp. 278. £35, \$70.

THE potential of photovoltaic solar-energy conversion as a renewable energy source is widely recognized. Photovoltaic power is already finding remote-site application in telecommunications, navigation aids and water pumping, for example, predominantly using crystalline silicon-cell technology. Calculators and watches with amorphous silicon and cadmium telluride solar cells are also now common. Surprisingly, there are few good general texts on the subject, which makes Van Overstraeten and Mertens's contribution all the more valuable.

The book begins with a comprehensive account of the background to the physics of solar cells, made particularly interesting by the inclusion of a thorough descrip-

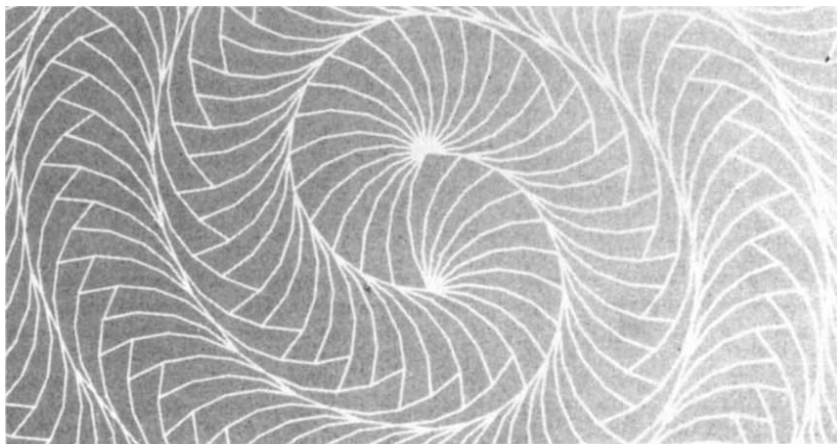
tion of the mechanisms which can account for loss of solar-cell efficiency. The theory is also developed separately for monocrystalline and polycrystalline silicon, amorphous silicon, gallium arsenide and other thin-film photovoltaic cells.

The method of manufacturing crystalline silicon photovoltaic modules is covered in great detail, from silicon purification and crystal growth, through cell fabrication, to the assembly of finished modules. In contrast, the section on thin-film technology is rather brief, with only a limited discussion of the various options. On the practical side, there are sections on measurement of the efficiency of light conversion of cells and modules, and their ability to withstand different environmental conditions, on cell interconnection in modules and on important factors in the design of photovoltaic systems for particular applications.

The book concludes with a summary of existing applications of photovoltaic systems which serves well to demonstrate the current dominance of crystalline silicon. This may well justify the authors' rather scant treatment of the emerging, potentially cheap, thin-film technologies such as amorphous silicon, cadmium telluride and copper indium diselenide cells which have yet to find large-scale commercial use beyond consumer applications.

Well written and logically structured, the book is easy to follow, and provides both a good text for the student and a useful reference work for the researcher. It describes the physics, technology and uses of photovoltaics for an audience anxiously awaiting a breakthrough in the availability of a low-cost, renewable energy source — we must hope that we don't have to wait too long. □

A. K. Turner is in the New Technology Division, British Petroleum Research Centre, Sunbury-on-Thames, Middlesex TW16 7LN, UK.



Singular construction — monohedral tiling (in which the tiles are all of the same shape) arranged into the semblance of a spiral; the prototile also admits many other, similar arrangements. The picture is reproduced from Tilings and Patterns, an unusual and happy marriage of mathematics and aesthetics written by Branko Grünbaum and G. C. Shephard and published by W. H. Freeman. The book will be reviewed in a future issue of Nature.

The many stages of carcinogenesis

D.G. Harnden

Introduction to the Cellular and Molecular Biology of Cancer. Edited by L.M. Franks and N. Teich. *Oxford University Press*: 1986. Pp.458. Hbk £30, \$59; pbk £15, \$32.50.

WILL Shakespeare started it: "One man in his time plays many parts". Since then analogies have run rampant. Now Beverly Griffin sees the seven ages of DNA — she does it well, and optimistically forecasts that in the seventh age the intricacies of the normal cell will be unravelled and lead to an understanding of cancer.

Perhaps I should try it — the seven ages of making a cancer textbook. First, the student, perhaps not "mewling and puking", but certainly confused by a subject such as oncology, where the range of disciplines is wide and information has to be picked up from many teachers and often from a variety of courses. Next, the teacher dedicated to a coordinated oncology course. Anyone who has tried, knows the problems. How can you even begin on DNA manipulation unless there is a firm grounding in DNA structure? Yet if you do not take much as read, there is not time to get from DNA to cancer therapy. And then, the brave editor, suggesting that the course should be converted into a book — the students are proffered more detail and time can be devoted to discussions. So distinguished experts send in their contributions, which somehow must be woven into an accurate and readable whole.

Much to their credit, Franks and Teich have succeeded to a large degree where others have totally failed. Their book will be of great value, not only to students but to anyone who wants a broad view of current knowledge of the scientific basis of cancer. The authors are indeed distinguished, the coverage is excellent and the prose very readable.

The groundwork is well laid by Franks himself, while chapters on metastasis and epidemiology ensure that the behaviour of cancer and the scale of the problem are well understood. But the heart of the book deals with carcinogenesis, cellular and molecular processes, and the nature of the cancer cell. The theme of stages set by Griffin recurs throughout the book. Multi-stage carcinogenesis is universally agreed, but each author sets these stages rather differently. For example, Adam sees radiation carcinogenesis in stages from the physical (10^{-12} to 10^{-18} seconds) to the cell and tissue stage which lasts months or years. Even growth factors operate in stages, and Waterfield spells them out

from the external signal to the nuclear response. Though the details of clinical cancer are beyond the scope of the work, the link to clinical problems is handled briefly but clearly.

This book is therefore at about stage four in its evolution; very good but not yet excellent. The authors, not too "jealous in honour", invite comment for its future improvement. Perhaps they could look to the order of the chapters. For example, why does Greaves's excellent chapter on the biology of leukaemia come so early? It makes no concessions to the uninformed and confronts complex ideas about cell lineages, maturation arrest and express-

ion of cell-surface molecules. The editors should also consider the difficult problem of how much background to fill in; some of the contributors expect sophistication of their readers, while others go back to first principles.

So onward to stage five. After that, stage six will be superb, "full of wise saws and modern instances". But what of stage seven? Let me predict — sans page, sans printed word, sans everything, except perhaps a computer terminal. □

D.G. Harnden is Director of the Paterson Institute for Cancer Research, Christie Hospital, Wilmslow Road, Manchester M20 9BX, UK.

Death on the Nile

D. M. Dixon

Science in Egyptology: Proceedings of the 'Science in Egyptology' Symposia. Edited by A. Rosalie David. *Manchester University Press/Longwood Press*: 1986. Pp.525 £45, \$77.50.

IN THE COURSE of excavations in Egypt and Nubia, carried out over several decades about the turn of the century, thousands of mummified and naturally preserved human bodies were recovered, ranging in date from pre-Dynastic times to the Christian period. Unfortunately, because of lack of specialist personnel on the staff of expeditions, the opportunities presented by this material were not fully exploited. With a few notable exceptions, excavators often simply selected for retention and later examination at home those bones which appeared to them to exhibit some abnormality or other point of interest. In general, the remains that did find their way to museums were for the most part neglected for years, during which specimens deteriorated and labels became detached.

Co-operation on any scale between Egyptology, science and medicine was slow in coming, and it was not until the mid-1960s that practical steps were taken to remedy the situation. The appearance of *Diseases in Antiquity*, edited by D. Brothwell and A.T. Sandison, and published in 1967 by Charles C. Thomas, marked a turning-point, although not all of the papers were concerned with Egypt. Shortly afterwards the work of Chiarelli and his colleagues at Turin, and the foundation of the Paleopathology Association in 1973 through the efforts of Aidan and Eve Cockburn, gave added impetus to interdisciplinary research on life in the ancient Nile Valley.

The Manchester Symposia of 1979 and 1984 have continued and expanded that work, the scope of the latter gathering being widened to include the discussion of

techniques applied to pottery and other antiquities. This volume contains a printed record of all the papers presented in 1984, together with a selection of those from 1979. Inevitably, the contributions vary considerably in value, but there is much of interest here. Among the most promising developments is the extraction and cloning in bacteria of DNA from mummified tissue, discussed by S. Pääbo, and the possibilities this offers of determining, among other questions, the genetic relationships between individuals and populations. This is clearly a matter of some importance to scholars grappling, for example, with the complexities of the history of Egypt during the latter part of the Eighteenth Dynasty.

An obstacle in the way of examining mummies has hitherto been the understandable reluctance of museum curators to allow the use of techniques which involve the total or partial destruction of specimens in their care. The application of endoscopy, described by E. Tapp and his colleagues, has changed things almost overnight. It has also permitted the taking of biopsies from such areas as the lungs and the confirmation of diagnoses based on radiological examination.

While the range of techniques discussed in this book is impressive, and augurs well for research on the lives of ancient peoples generally, a note of caution is necessary. There is still a long way to go before Egyptologists can turn to their medical and scientific colleagues with the reasonable expectation of receiving satisfactory answers to their problems. In this regard Renate Germer's closing paper, "Problems of Science in Egyptology" makes salutary reading. As she reminds us, much of the information obtained through the use of methods discussed at the Symposia is at present of limited help to Egyptologists because "a great deal of basic research has yet to be done to determine the applicability of the scientific methods themselves". □

D. M. Dixon is a Lecturer in the Department of Egyptology, University College London, Gower Street, London WC1E 6BT, UK.

Glittering prizes for research support

David F. Horrobin

Could public support for research be cheapened, and made more productive, by following the eighteenth-century precedent of the British government's prize for a means of measuring longitude?

THE crisis in British universities, the shortages of research funds and the insistent demands that science should serve the economic needs of the nation which supports it, ought to require a fundamental rethinking of our approach to the financing of science. I see much evidence of panicky tinkering by both the scientific community and the government, but few indications of radical thought.

We should perhaps consider what may be the first example of parliamentary funding of research, which is also one of the most successful. Because it is such an excellent precedent, those in power might take it seriously. The hero is Mr John Harrison, born in 1693 the son of a poor Yorkshire carpenter, a self-educated clockmaker and the toast of Neil Armstrong at a twentieth century 10 Downing Street dinner.

Petition

In a petition to the British Parliament on 25 March 1714, Captains of Her Majesty's Ships, Merchants of London and Commanders of Merchant Men demanded a solution to the problem of measuring longitude at sea. Failure was causing innumerable navigational disasters with the loss of ships, men, goods and battles. The government, wisely, chose not to throw money at the problem by supporting research by experts who, by definition, had failed to solve it. Instead, they arrived at an astonishing proposal: a prize of £10,000 was offered for a method that would reliably measure longitude to within 1°; £15,000 for a method accurate to within 40'; and £20,000 for a method accurate to within 0.5°. In 1714 those sums were truly astronomical. Anyone winning the prize would become not merely comfortable but genuinely rich.

The result was an outburst of private research on longitude. The problem and the prize were known to everyone. They are mentioned in Swift's *Gulliver's Travels* and a Hogarth cartoon shows a lunatic obsessed with them. Theoretically, it was known that one way to solve the problem was to make a timepiece so accurate that it would tell the time at Greenwich even after many weeks at sea, much buffeting and drastic changes in temperature. Comparison of that time with local time would then allow an accurate calculation of longitude.

None of the experts believed that a timepiece could be so accurate. An array

of bizarre and complex proposals was put forward by the Fellows of the Royal Society and their friends, but all failed. Only poor, provincial, self-taught John Harrison believed that a clock could be made to the required standards. His story is well known: his production of increasingly reliable, robust and accurate chronometers, their objective success in Royal Navy trials, his denigration by the eminent gentlemen of the Royal Society, the attempts by the Society's experts to deny him the prize legitimately won, the anger of the King when he heard about these academic machinations and Harrison's final triumph with grant of the full £20,000 at the age of 80. But the lessons to be drawn from this remarkable story are as valid now as they were in the eighteenth century.

- The announcement of a cash award sufficiently large to make the winner enormously wealthy will generate a phenomenal outpouring of research directed to the solution of practical problems. Such research will be privately financed and will cost the donor nothing but the cost of the award.
- Such an award will lead to unprecedented cross-fertilization of fields of endeavour. If the prize is large enough, scientists whose speciality may appear remote from the problem will be stimulated to think about it.
- The definition of success must be crystal clear, practical and open to verification by non-experts. The solution need not be brilliant or sophisticated, nor need it meet with the intellectual approval of experts. The only condition is that it must work.
- There must be no limitation on the categories of people allowed to succeed. The challenge must be open to non-experts and to the sons of poor carpenters.

Prizes

The government should decide what problems it wants solving. In my field of medicine, obvious problems whose solutions would save a great deal of money include schizophrenia, eczema, multiple sclerosis and Alzheimer's disease. People from other fields should be able to produce long and comparable lists. Economists could then work out what each particular problem costs the nation, and a prize could be offered for a practical approach which would either eliminate the problem or reduce the cost of solution. The prize should

be the value of the savings made during one full year.

Some of the prizes might turn out to be very large indeed, tens or even hundreds of millions of pounds. Even greater amounts would be possible if, say, the European Community, the United States and other countries acted together, as they might well do if the scheme cost little or nothing. The prize should be given tax free to the individual, company or group producing an answer.

The advantages of the scheme would be many, but include:

- The cost of initiating the scheme would be trivial. After the first year, a successful solution would save millions.
- Researchers of all kinds would be encouraged but not forced to work on practical problems. They would also be encouraged to extend their range of expertise.
- There would be large inflows of private venture capital into research as individuals, companies and others chose to support research directed towards the solution of problems attracting the larger prizes.
- There would probably be rapid progress in basic science. Many now suffer from the illusion that basic discoveries always precede practical research. Historical evidence suggests that the traffic is just as frequently in the other direction. We have forgotten the story of Pasteur. Brilliant people devoting themselves to practical problems, such as the spoiling of wine or dying silkworms, may end up by discovering new fundamental principles.
- The enterprise would be cash-generating rather than cash-consuming. Governments would be delighted with science instead of looking at it with a jaundiced eye.
- Last, but by no means least, the huge cash prizes would put money into the hands of the most practically creative members of society. The money would certainly be spent in interesting ways; on further research, on enterprising investment or even on innovative techniques of personal consumption.

The idea would cost almost nothing to put into operation. It would rejuvenate science and transform attitudes to science. It should be given a try. □

David F. Horrobin is managing director of Efamol Ltd, Woodbridge Meadows, Guildford GU1 1BA, UK.

Uncertainties in the smoke source term for 'nuclear winter' studies

Joyce E. Penner

Lawrence Livermore National Laboratory, University of California, Livermore, California 94550, USA

Climate models have shown that the effects on climate of a major nuclear exchange depend on the quantity and optical properties of the smoke that is dispersed into the global atmosphere. Published estimates for each of the factors necessary to determine the smoke optical depth yield a wide range of values, which allows the prediction of either comparatively minor effects on climate or massive effects.

THE potential effects on climate of large amounts of smoke injected into the atmosphere following a major nuclear exchange have been widely analysed¹⁻⁴. Although simplifications and uncertainties still exist in the application of climate models to calculate the effects of smoke, many of the simplifications that had to be made in the first studies have now been corrected. These improved climate models have shown that the effects of smoke on climate depend on the quantity and optical properties of the smoke that is generated and dispersed into the global atmosphere.

The amount of smoke and its optical properties can be summarized by the average optical depth that would result if the smoke were dispersed throughout half the Northern Hemisphere. If the extinction optical depth $\bar{\tau}_e$ is taken as the sum of the scattering and absorption optical depths, it may be calculated from $\bar{\tau}_e = (k_s + k_a)S/A$, where k_s and k_a are the scattering and absorption cross sections of the smoke (in $\text{m}^2 \text{g}^{-1}$), S is the amount of smoke (in g) and A is the area blocked by the smoke cloud (taken as half the area in the Northern Hemisphere, or $1.28 \times 10^{14} \text{m}^2$). The quantity of smoke may be calculated from $S = \epsilon F(1 - f_r)$, where ϵ is the emission factor of smoke (g smoke per g of fuel burned), F is the amount of fuel burned, and f_r is the fraction of smoke removed by precipitation in the convection column above the fires and in the first few days after the war, during which the smoke is presumed to spread out to global scales. The original 'baseline' analysis of Turco *et al.*² resulted in an estimate for $\bar{\tau}_e$ of ~ 6 for smoke produced in urban fires. Their smoke had a single scattering albedo of ~ 0.5 , so that $\bar{\tau}_a = k_a S/A$ was ~ 3.0 . Consideration of the additional smoke from wildlands fires, long-term fires and dust from surface bursts increased the estimate for $\bar{\tau}_e$ to ~ 8 . (Turco *et al.*² spread their smoke over the entire Northern Hemisphere, so that their published optical depths differ from these by a factor of 2.) All subsequent analyses have similarly implied 'best estimates' for $\bar{\tau}_e$ that were substantially greater than 1. Here I review the estimates for the various factors which contribute to $\bar{\tau}_e$ and $\bar{\tau}_a$, to obtain reasonable bounds on the range of magnitudes consistent with current knowledge. This range encompasses values that may be associated with major climatic effects if a large fraction of the available urban combustible load is burned. On the other hand, within the bounds set by current analysis, comparatively minor effects are also possible, especially if the targeting of weapons avoids refineries and other large storage facilities that contain petroleum or other fossil fuels. I then recommend several areas for research that could lead to more certain estimates of the effects of such a war.

Amount of fuel burned in urban fires

As pointed out in ref. 5, several methods have been adopted for estimating the amount of fuel that might burn in urban fires. These methods are not mutually consistent. In the method

adopted by Turco *et al.*², which yields the highest fuel estimate, the amount of fuel burned in urban fires is determined from the product $F = \text{FL} \times f_b \times A_i \times \text{SF}$ where FL is the average areal fuel load within the burned-out region (in g m^{-2}), f_b is the fraction of fuel that is consumed by fire within the area that burns, A_i is the area that is initially ignited by the fireball (in m^2) and SF is the average areal spread factor for the fires.

An overestimate by this method may be caused by at least two factors. Most often, no account is taken of the overlap of burned areas when detonations take place near one another. Secondly, the entire ignited area and its radially expanded spread is assumed to coincide exactly with the urban fuel bed and with the average fuel load FL. This assumption may be seriously in error, for example, for targets such as airports that generally reside on the outer edges of cities. It is often argued that these effects are mitigated by the choice of a 'conservative' value for A_i —that is, one that corresponds to the area that would be ignited by a thermal fluence of 20 cal cm^{-2} , rather than the area associated with a thermal fluence of $7\text{--}10 \text{ cal cm}^{-2}$, which is considered sufficient to ignite at least the lighter fuel elements such as paper and twigs. In Nagasaki, where the presence of hills restricted the fire ignition area⁶, the actual area burned ($A_i \times \text{SF}$) corresponded to the area which would have received a thermal fluence of 20 cal cm^{-2} ; whereas in Hiroshima, the area burned corresponded to the area which received only 7 cal cm^{-2} . It would therefore seem that the area corresponding to 20 cal cm^{-2} is indeed 'conservative', if $\text{SF} = 1$.

It is not known, however, whether the two effects mentioned above (overlap and improper average values for FL) would indeed be balanced by an underestimate for $A_i \times \text{SF}$. Several lines of inquiry suggest that the overestimate may be significantly larger than the factor of 2 underestimate made by using a fluence area corresponding to 20 rather than $7\text{--}10 \text{ cal cm}^{-2}$. The analysis in ref. 7, which consisted of a large attack on cities, suggests that consideration of overlap may reduce the value of A_i by as

Table 1 Average combustible fuel load in cities

Authors	Fuel load (kg m^{-2})
Turco <i>et al.</i> ²	
Baseline case	33.5*
100-Mtonne, city centre	200.0
Crutzen <i>et al.</i> ¹⁰	40.0†
NRC ³	40.0
Reitter <i>et al.</i> ⁸	
Detroit, centre	34.5
Detroit/San Jose, suburbs	10.2

* Average of 100 kg m^{-2} in 'city centres' and 30 kg m^{-2} in suburbs.

† This estimate was not used in the final analysis of ref. 10 (see text).

much as a factor of 4. Also, overlap of ignition areas may reduce fire spread, if intense firestorms are generated which create inflowing winds. On the other hand, with more intense fires, the fuel may burn more thoroughly, increasing f_b .

The value used for the average fuel load can be checked by an analysis of some of the fire spread modelling results of Reitter *et al.*⁸ (see also Appendix 3A of ref. 4). All previous studies of the amount of fuel that might burn have assumed average values for FL of $\sim 40 \text{ kg m}^{-2}$ (see Table 1), although values from 10 to 400 kg m^{-2} are quoted as possible³. Table 1 also shows the average areal fuel loads within the ignited areas corresponding to a 1-Mtonne nuclear explosion over the centre of Detroit and several 1- and 0.5-Mtonne bursts over detonation points above suburban Detroit and central San Jose, as taken from Reitter *et al.*⁸. These data imply that values closer to 10 kg m^{-2} should be used for most urban and suburban areas. Furthermore, if the fuel loads for Detroit are correct, values for FL of $\sim 40 \text{ kg m}^{-2}$ are appropriate only for weapons directed at the centres of large cities. The average fuel loads from Reitter *et al.* consider only those areas occupied by buildings, so that these average values do not account for any decrease in FL due to targeting on the fringes of cities or near lakes or parks which would have lower average fuel loads. The fuel loads were developed from surveys taken in the late 1960s, but recent analyses of fuel loads in San Jose⁹ are similar to the average fuel load used by Reitter *et al.* for that city. The Detroit fuel loads used by Reitter *et al.* seem surprisingly low, but result from averaging the fuel loads in the centre of that city (which range up to 160 kg m^{-2} over small areas) with values near 10 kg m^{-2} in outlying regions. Note that Detroit has a population of >4 million. There are only 39 urban centres in the world with a population of >3 million, and only 80 cities in the NATO and Warsaw Pact countries with populations of >1 million. The oft-quoted "100-Mtonne central city" model of Turco *et al.*², assumed values for FL of 200 kg m^{-2} in 100 cities: these loads appear to be overestimated by a factor of 5. Of course, one must check whether European cities or cities in the eastern United States contain much higher fuel loads than those quoted here for Detroit. But it seems highly probable that the estimates for FL used previously are too large, especially in view of the analysis of total fuel load outlined below. Crutzen *et al.*¹⁰ also rejected this method in favour of the inventory approach outlined below. Finally, note that the above analysis for FL assumed targets which were entirely contained within the urbanized areas occupied by buildings. Consideration of actual target locations, some of which will occur on the fringes of cities and some of which will fall near lakes or other low-fuel-density areas, will further reduce the estimate for FL.

Significant further reduction of the uncertainties using this approach requires a detailed analysis on a city-by-city basis, with consideration of specific target locations, fuel loads and overlap of fire areas. Here we consider an alternative approach, in which total combustibles are estimated directly and then a fraction is assumed to be ignited and burned.

The total inventory of combustibles has been estimated by Crutzen *et al.*¹⁰ and by Bing⁵, using different methodologies. Pittock *et al.*⁴ also estimated inventories, but as ref. 4 is primarily a review of ref. 10, they will not be separately quoted here. Crutzen *et al.*¹⁰ used production figures for various raw materials and estimates of their lifetimes to obtain estimates for the total abundance of cellulosic materials, polymeric materials and asphalt. Bing⁵, on the other hand, gathered data from surveys of fuel loads in various types of structures and their contents for the United States, and extrapolated these data to Europe and the Soviet Union. The two sets of published figures are not directly comparable because Crutzen *et al.* estimate the amount of cellulosic and polymeric materials in the developed world, whereas Bing's estimates refer only to the NATO and Warsaw Pact countries. Crutzen *et al.* and Bing also separately estimate the amount of petroleum available to burn, including petroleum

stored as primary stocks and as secondary stocks. The figures of Crutzen *et al.* refer to the amount of petroleum stored globally, whereas Bing's numbers again refer only to that fraction contained within the NATO and Warsaw Pact countries. In order to consider similar models, I have reduced the inventories published by Crutzen *et al.* for the developed world by the ratio of the population of the NATO and Warsaw Pact countries to that of the developed world. Their estimates for petroleum were similarly reduced by the ratio of consumption rates in NATO and Warsaw Pact countries to that of the world. As shown in Table 2, Crutzen *et al.*'s inventory still implies 2.5 times more cellulosic material than does Bing's. Remarkably, as the methodologies were different, the estimates for petroleum-derived materials are similar. The total amount of fuel in primary stocks of petroleum is fairly well known and the two estimates are comparable. The amount of fuel in secondary stocks of petroleum is less well known. Two different sources of data have led to a difference of a factor of 2–4 for fuel in this category, as shown in Table 2.

The methodologies used in these two studies have obvious difficulties, and it is not clear which method is more appropriate. Note, however, that the totals for cellulosic and polymeric materials assumed by Crutzen *et al.*¹⁰ are not entirely consistent with the fuel load estimates derived from Reitter *et al.*⁸. For

Table 2 Inventory of total available combustibles in NATO and Warsaw Pact countries (Tg)

Authors	Cellulosic materials	Primary petroleum stocks	Secondary petroleum stocks	Polymeric materials
Crutzen <i>et al.</i> ¹⁰	16,500*	462†	198–462†	574*
Bing ⁵	6,444	480	100	753

* Reduced from the estimate for the 'developed world' in ref. 10 by the ratio of populations, 0.87 (G. Bing, personal communication).

† Reduced from the estimate for the whole world in refs 4 and 10 by the ratio of consumption rates in NATO and Warsaw Pact countries to that of the world, 0.66.

example, we may use the average areal fuel load for urban and suburban areas from ref. 8, together with the total urban area in cities with population greater than 2,500 in the United States, $135,000 \text{ km}^2$, to arrive at a combustible load for the United States of $1,350 \text{ Tg}$. Consideration of 50 city centres with fuel loads similar to that of Detroit might increase this total to $2,000 \text{ Tg}$. This number is close to the value derived by Bing⁵ for the United States ($2,119 \text{ Tg}$), and thus lends confidence to his estimates. On the other hand, when Crutzen's numbers for the developed world are scaled by the ratio of population in the United States to that in the developed world (0.225) we arrive at $4,400 \text{ Tg}$, which is at least twice as large as the estimate above. In the analysis below, both numbers will be used to estimate the range of optical depths that are possible, given current uncertainties.

Table 2 summarizes the inventories of combustibles in NATO and Warsaw Pact countries, derived by these two methodologies. To allow consideration of the range of smoke absorption properties from various fuel types, Table 2 divides the inventories into cellulosic fuels, petroleum-derived fuels and liquid fossil fuels. This last category has been subdivided into primary and secondary stocks of petroleum. Secondary stocks are considered to be distributed with other fuels, whereas primary stocks of petroleum are considered separately in order to calculate the effect of a concerted effort to avoid or include these targets (see estimates below).

Typically, only a fraction of the total available combustible material might actually burn in active, flaming combustion. Here I assume this fraction to be 25% of the distributed fuels (cellulosic and polymeric materials, and secondary stocks of petroleum) for both the high and low estimates. In this way the range of optical depth derived can be considered to be independent of any particular model, although more (or less) fuel might

burn if the warring nations made a concerted effort to try to ignite (or avoid burning) the available fuel. A 25% fraction might come about, for example, by associating ~65% of the total fuel with people who live in cities (the average proportion of city dwellers for Europe, the Soviet Union and the United States), and then burning 80% of the total fuel in cities, half in active, flaming combustion and half in longer-term, smoldering combustion. Alternatively, most city fuels might burn in active combustion but, because of clustering of targets and overlap of ignition areas, only 40% of the fuel in cities actually ignites and burns. The optical depths derived below could easily increase if the war involved, for example, attacks on oil rigs or on oil-producing countries whose fuel stores are not part of the inventory considered here. Similarly, if attacks included population centres outside the countries included here, the optical depths would increase. However, the NATO and Warsaw Pact countries contain 85% of the population of all countries in the developed world, and as the present model burns 40–80% of the urban fuels in these countries, it is certainly large enough to be considered a 'major' nuclear war. On the other hand, ~30% of the population in the developed world is concentrated in the 100 or so largest cities (see Table 3.1 in ref. 4), so that if 100% of the fuel of these cities is burned in active, flaming

Table 3 Smoke emission factor

Authors	Per cent of fuel
Turco <i>et al.</i> ²	2.7*
Crutzen <i>et al.</i> ¹⁰	
Wood	1.5
Oil, polymers, etc.	7.0
NRC ³	
Wood	3.0
Oil, polymers, etc.	6.0
Range of values in flaming combustion ¹⁰	
Wood	0.085–2.5
Oil	2–10
Plastics	1.2–50

* Weighted average of the 'net emission factors' of 1.1% for urban centres and 3.3% for suburbs. These values may include some allowance for scavenging by rain. Emission factors in an earlier version of ref. 2 were 2.5% for city centres and 5% for suburbs.

combustion, nearly the same total optical depth could be produced. However, the purpose here is not to develop war models, but to show the range in optical depth expected from current knowledge for a given model. These ranges can easily be adapted to other models.

Smoke emission factor

The appropriate smoke emission factor in a large-area urban fire depends on a number of poorly estimated and poorly known factors. Emission rates can vary^{11,12}, depending on the type of fuel, the ambient air temperature, the availability of oxygen, the radiant intensity (as determined by the proximity of nearby fires), the geometric arrangement of fuel, and so on. Only very limited data from large fires are available, so most studies have used values consistent with the range of emission factors measured in laboratory-scale fires (see Table 3). These might be underestimates, if oxygen availability is truly limited in a large-area fire. On the other hand, Carrier *et al.*¹³ have argued that oxygen availability should not be an issue, given the turbulent motions above the fire. In view of the lack of credible data for smoke emission factors from large-area fires, the values adopted here are estimated from the limited available data, and are highly uncertain. Table 3 also includes the range of estimates of the emission factor compiled in ref. 10, primarily from labor-

Table 4 Absorption and extinction coefficients for smoke from urban fires

Authors	k_a ($m^2 g^{-1}$)	k_e ($m^2 g^{-1}$)
Turco <i>et al.</i> ²	2.9	5.8
Crutzen <i>et al.</i> ¹⁰		
Wood	3.3	6.8
Oil, etc.	7.0	10.5
NRC ³	2.0	5.5
Penner and Porph ¹⁵		
Wood, no coagulation	1.5	6.6
Oil, etc., no coagulation	5.6	9.5
Wood, after coagulation	1.3	4.0
Oil, etc., after coagulation	1.8	4.0

atory data. In the analysis below, I adopt a range of values for the emission factor, consistent with the best estimates chosen in refs 3 and 10; however, larger uncertainties apply because of the possible inapplicability of these emission factors to large-area fires.

Optical properties of smoke

Just as the emission factor for smoke depends on the burning conditions and type of fuel, so does the chemical, morphological and optical character of the smoke. Nevertheless, various authors have estimated the absorption and extinction coefficients for smoke, based on a variety of measurements that have tended to emphasize the data available from smoke emitted under flaming conditions. The estimates are summarized in Table 4. In most recent evaluations, the optical properties of wood smoke are distinguished from those of smoke from fuels such as oil, plastics, and other polymers whose chemical structure has little available oxygen. These latter fuels tend to produce much blacker smoke. Table 4 shows wide variations in the estimates of the absorption and extinction coefficients for fresh smoke.

In addition to variations in average optical absorption and extinction from different evaluations of these properties for fresh smoke, two mechanisms may act to make aged smoke less absorbing (see ref. 4). The first mechanism is coagulation, which may act on short timescales (in very dense smoke plumes) or on longer timescales (from days to a week in the spreading global plume) to create larger particles. These larger particles would absorb and scatter less radiation, if they are spherical. Because some smoke particles are quite oily (and therefore spherical), while others appear as fluffy or chained agglomerates, it is not possible to predict the effects of coagulation on optical properties. Chained agglomerate particles might become spherical if they coagulated with oily smoke particles, or if they were to condense and re-evaporate in a cloud—a process which might allow the chains to collapse¹⁴. To the extent that the agglomerates remain in a chained formation, their absorption properties may not change significantly. Thus, in the following, I adopt two extremes. In the first case, coagulation is assumed to have no effect on optical properties; in the second, coagulation is assumed to reduce extinction and absorption by the amounts estimated in ref. 15 for several days of coagulation. This additional consideration widens the discrepancy between the lowest and highest estimates of absorption coefficient by an additional factor of ~3 for the highly carbonaceous, absorbing smokes. The extinction coefficients differ by a factor of >2. The absorption coefficient for less absorbing smoke is not significantly changed by coagulation.

Fraction of smoke scavenged by rain

The last factor which contributes to estimates of the average optical depth is the amount of smoke that is removed by precipitation occurring in the smoke plume over the fire and in the

Table 5 Fraction of smoke removed by rain

Authors	Fraction removed
Turco <i>et al.</i> ²	
Suburban fires	0.25
Firestorms	0.50
Crutzen <i>et al.</i> ¹⁰	0.30
NRC ³	0.50
Cotton ¹⁷	~0.02*
Penner and Edwards ¹⁸	0.50†

* Fraction of smoke particles that enter cloud water by phoretic scavenging. The amount removed by precipitation may be less than this.

† Fraction of smoke particles that enter cloud water by nucleation scavenging. The amount removed by precipitation may be less than this.

first few days after the war. Several authors have estimated that, especially for large, intense fires, large quantities of water will condense above the fire^{16,17}, but there have been few attempts to quantify how much smoke is removed. In both Hiroshima and Nagasaki, a 'black rain' fell, coincident with the fires which followed the nuclear blasts of August 1945⁶. The black rain is presumably smoke that has been scavenged by rain. It is not known how much of the smoke was scavenged or whether the experience in Nagasaki and Hiroshima should be considered to be typical.

The amount of smoke scavenged by rain depends, once again, on properties of the smoke which are poorly known. For example, the number of smoke particles that act as condensation nuclei for cloud drops depends on the highest level of supersaturation attained above the fire. This depends on the total number of smoke and debris particles, their size distribution, and their affinity for water. Furthermore, the highest level of supersaturation depends on the updraft velocity within the plume as well as the growth rate of the drops which form. As the size and composition of smoke particles from the large-scale urban fire are poorly known, it is difficult to predict how much nucleation scavenging might occur. Initial studies, which assume idealized spherical particles, indicate that for conditions considered typical of smoke above fires, all particles with radii greater than 0.1 μm (or ~50% of a log-normal distribution with a mode radius of 0.1 μm) might be scavenged in this manner¹⁸. However, further studies with a more realistic description of smoke particles are needed.

Once drops have formed, other mechanisms may also act to

attach smoke particles to cloud drops. These include electrical capture, phoretic forces and turbulent motions, although Cotton¹⁷ has estimated that <2% of the smoke particles might be captured phoretically. Cloud drops may or may not form precipitation-sized rain drops. The probability of this occurring depends on the initial size of debris and smoke particles and on the number that become nucleated to form drops. Pruppacher¹⁹ estimated that, although most of the particles would enter cloud drops by nucleation scavenging, rainout of the smoke would be relatively small because of overseeding effects. However, only one size distribution was considered and the updraft velocity was only 1 m s^{-1} . Consideration of more realistic fire plume conditions may alter this conclusion. Once the drops become large enough to obtain a significant fall velocity, they may capture more smoke particles by impaction scavenging. The probability of this occurring also depends on the size of smoke particle (with larger particles being more likely to be scavenged).

The capture mechanisms described above apply to warm-rain precipitation only. Additional mechanisms and pathways for capture must be considered in the case of ice formation, which can also lead to release of the smoke particles. Because the saturation vapour pressure over ice is less than that over water, when ice begins to form, water on drops may evaporate and recondense on ice particles, thereby releasing any smoke that was previously captured by nucleation scavenging. Recently G. Tripoli (personal communication) has used the cloud model developed by Cotton¹⁷ to estimate that this process could limit any rainout of the smoke above the fire to <20%. In addition to the near-immediate rainout of smoke above the fire, mesoscale atmospheric circulations may be set up which lead to rainout on longer timescales²⁰.

The theoretical analysis of scavenging and rainout is complex and difficult, and further work is needed. For this reason, many authors have simply guessed a fraction for smoke that might be removed by rainout. These guesses range from ~0 to 50% (see Table 5), although the real range of possibilities might include values up to 100% in some cases²¹. Here I consider the range from 0 to 50%. Obviously, the range of average optical depths obtained could be larger, for example, if rainout removed 90% of the smoke.

Calculated optical depths and climate

If we combine all the choices described above, emphasizing the smallest factors in one case and the largest factors in the second,

Table 6 Calculated range of average optical depth from urban fires

Category	k_a ($\text{m}^2 \text{g}^{-1}$)	k_c ($\text{m}^2 \text{g}^{-1}$)	ϵ (g g^{-1})	F (Tg)	$(1-f_r)$	$\bar{\tau}_a$	$\bar{\tau}_c$
Wood							
High	3.3*	6.8*	0.03†	16,500*/4	1.0‡	3.19	6.57
Low	1.3(1.5)§	4.0(6.6)§	0.015*	6,444 /4	0.5†	0.12(0.14)	0.38(0.62)
Polymers, plastics, etc., and secondary stocks of petroleum							
High	7.0*	10.5*	0.07*	1,083¶/4	1.0‡	1.04	1.55
Low	1.8(5.6)§	4.0(9.5)§	0.06†	674#//4	0.5†	0.07(0.22)	0.16(0.38)
Primary stocks of petroleum							
High	7.0*	10.5*	0.07*	480	1.0‡	1.84	2.76
Low	1.8 (5.6)§	4.0 (9.5)§	0.06†	462*	0.5†	0.19 (0.61)	0.43 (1.03)

* From refs 4 and 10.

† From ref. 3.

‡ Based on ref. 17.

§ From ref. 15. Numbers in parentheses refer to the case with no coagulation.

|| From ref. 5.

¶ This number is the sum of the average of high and low estimates for secondary stocks of petroleum from ref. 10 and Bing's⁵ estimate for polymeric materials (see Table 2).

This number is the sum of Bing's⁵ estimate for secondary stocks of petroleum and Crutzen *et al.*'s¹⁰ estimate for polymeric materials (see Table 2).

Table 7 Average optical depth

Model	$\bar{\tau}_a$	$\bar{\tau}_e$
Distributed fuels only		
High	4.2*	8.1
Low	0.2†(0.4)‡	0.5(1.0)
With primary stocks of petroleum		
High	6.1§	10.9
Low	0.4 (1.0)	1.0(2.0)

* Equivalent to 270 Tg of smoke with $k_a = 2 \text{ m}^2 \text{ g}^{-1}$.

† Equivalent to 12 Tg of smoke with $k_a = 2 \text{ m}^2 \text{ g}^{-1}$.

‡ Optical depths in parentheses refer to the case with no coagulation.

§ Equivalent to 388 Tg of smoke with $k_a = 2 \text{ m}^2 \text{ g}^{-1}$.

|| Equivalent to 24 Tg of smoke with $k_a = 2 \text{ m}^2 \text{ g}^{-1}$.

we obtain the range in absorption and extinction optical depths shown in Table 6. Table 6 lists separately the optical depths from cellulosic fuels, from distributed fuels producing highly carbonaceous smoke (polymeric materials and secondary stocks of petroleum) and from primary stocks of petroleum. In the case of distributed fuels, one quarter of the total abundance in the NATO and Warsaw Pact countries, or 40–80% of the fuel in cities, is assumed to burn. As shown in Table 6, the burning of polymers and petroleum may contribute significantly to the total optical depth. To show its effect, we consider two models. In the first, the contribution of primary stocks of petroleum to the total optical depth is not included. This might result if the warring nations specifically tried to avoid targets, such as refineries, that would add disproportionately to the optical depth. As warring nations have never yet refrained from attacking the economic base of their enemies, this may be an unlikely model; but it is interesting to calculate the consequences should studies of 'nuclear winter' cause nations to re-evaluate their war plans. In the second model these targets are all included, so that 100% of the primary stocks of petroleum are burned. Table 7 summarizes the high and low estimates of optical depth for these two cases. If the primary stocks of petroleum are not included, the absorption optical depth varies from 0.2 to 4.2. Including these stocks increases the range of absorption optical depths to 0.4–6.1. In the first case, the low estimate is equivalent to 12 Tg of smoke with the optical properties assumed in ref. 3. This increases to 24 Tg of smoke if primary stocks of petroleum are included. This case is close to the lowest amount of smoke (20 Tg) considered by Malone *et al.*²² in an advanced three-dimensional climate simulation. Their results are consistent with widespread temperature changes of -4 to -6°C over the continents in summer. The largest average optical depth calculated here is equivalent to almost 400 Tg of smoke (assuming the absorption coefficient from ref. 3). This is somewhat less than the largest amount of smoke assumed by Malone *et al.*²² (500 Tg)

and would, according to their results, lead to profound climate changes, particularly as its removal would be inhibited by changes to atmospheric stability.

The range of values calculated here is disquieting, because I tried to choose values for each of the various factors that were thought to be a 'best estimate' by at least one of the authors cited here. Although the range derived is similar to that quoted in ref. 3, they developed their estimates from uncertainties about a median value, whereas the present estimates are developed using the best estimates from a number of authors. The calculated range would be larger if we also considered the uncertainty estimates for each of the factors. This is especially true for the range of values chosen for the emission factors. In this case, the range of published values is not very large, but, because so few relevant data are available, the published range may not even include the correct value. There are also uncertainties caused by the lack of good data on the optical properties of smoke and the effect of the clouds formed above large fires. Furthermore, as discussed above, little is known about the scavenging and rainout of smoke in fire plumes. Good data on the properties of smoke from large fires will only come from large fire experiments; but great care is needed in the design and interpretation of the large-scale fires which will be used in the re-analysis of the smoke source term. The planned experimental programs sponsored by the United States Defense Nuclear Agency should be helpful but must not stop after only the first few experiments. We must try to understand the more complex situations that will exist in a real nuclear fire. In addition, more and greater emphasis must be placed on understanding scavenging and rainout. Here, progress may come through the development of advanced modelling capabilities, coupled with verification by large-scale fire experiments.

Conclusions

Although our lowest estimates for $\bar{\tau}_a$ and $\bar{\tau}_e$ may produce only minor climatic effects, models can easily be constructed in which more fuel is burned, so that even in the low-estimate case, the estimate for $\bar{\tau}_a$ would correspond to a major climatic impact if the war takes place in the spring or summer. On the other hand, it is entirely possible that such major effects could be avoided if the low estimates are correct and if targets such as refineries, oil and gas production fields and coal storage areas are avoided. The effects on climate could also be lessened if the war took place during the winter⁴. It seems clear, therefore, that 'nuclear winter' is not necessarily a probable outcome of nuclear war, although it is certainly possible. The full range of possible impacts can never be completely narrowed because we cannot have access to the war plans of the nations of the world, nor predict the course of any given war once it began. I have shown, however, that for the model considered here, that is, one in which 40–80% of the total distributed urban fuels are burned, further research is needed in order to be able to predict the effects on climate.

1. Crutzen, P. J. & Birks, J. W. *Ambio* **11**, 114–125 (1982).
2. Turco, R. P., Toon, O. B., Ackerman, T. P., Pollack, J. B. & Sagan, C. *Science* **222**, 1283–1292 (1983).
3. National Research Council *The Effects on the Atmosphere of a Major Nuclear Exchange* (National Academy Press, Washington, 1985).
4. Pittock, A. B. *et al. Environmental Consequences of Nuclear War Vol. 1* (Wiley, Chichester, 1986).
5. Bing, G. *Lawrence Livermore natn. Lab. Rep. No. UCRL-93192* (1985).
6. Ishikawa, E. & Swain, D. L. *Hiroshima and Nagasaki, The Physical, Medical, and Social Effects of the Atomic Bombings* (Basic Books, New York, 1981).
7. Levi, B. G. & Rothman, T. *Physics Today* **38**, 58–65 (1985).
8. Reiter, T. A., Kang, S.-W. & Takata, A. N. *Lawrence Livermore natn. Lab. Rep. No. UCRL-53647* (1985).
9. Simonett, D. S., Barrett, T. N., Gopal, S., Holzmüller, F. J. & Sun, G. preprint (University of California, Santa Barbara, 1986).
10. Crutzen, P. J., Galbally, I. E. & Brühl, C. *Clim. Change* **6**, 323–364 (1984).
11. Bankston, C. P., Cassanova, R. A., Powell, E. A. & Zinn, B. T. *U.S. natn. Bur. Standards Rep. No. NBS-GCR-78-147* (1978).
12. Twarson, A. preprint (Factory Mutual Research, Norwood, Massachusetts, 1984).
13. Carrier, G., Fendell, F. F. & Feldman, P. *Combust. Sci. Technol.* **39**, 135–162 (1984).
14. Goldsmith, P., May, F. G. & Wiffen, R. D. *Nature* **210**, 475–477 (1966).
15. Penner, J. E. & Porch, W. M. *Atmos. Envir.* (in the press).
16. Penner, J. E., Haselman, L. C. Jr & Edwards, L. L. *J. Clim. appl. Met.* **25**, 1434–1444 (1986).
17. Cotton, W. R. *Am. Scient.* **73**, 275–280 (1985).
18. Penner, J. E. & Edwards, L. L. in *Proc. Conf. Cloud Physics* **2**, 83–86 (Am. meteor. Soc., Massachusetts, 1986).
19. Pruppacher, H. R. in *Proc. int. Semin. Nuclear War* (eds Newman, W. S. & Stipcich, S.) 163–174 (Servizio Documentazione dei Lab. natn. di Frascati dell'INFN, Erice, 1985).
20. Golding, B. W., Goldsmith, P., Machin, N. A. & Slingo, A. *Nature* **319**, 301–303 (1986).
21. Hobbs, P. V., Radke, L. F. & Hegg, D. A. in *Proc. ICSU-SCOPE Workshop Nuclear Winter Scenario: Current Studies of Scavenging Models* (International Committee on Cloud Physics, Tallinn, Estonia, in the press).
22. Malone, R. L., Aver, L. H., Glatzmaier, G. A., Wood, M. C. & Toon, O. B. *J. geophys. Res.* **91**, 1039–1053 (1986).

Occultation determination of Neptune's oblateness and stratospheric methane mixing ratio

E. Lellouch*, W. B. Hubbard†, B. Sicardy*, F. Vilas§ & P. Bouchet||

* Observatoire de Meudon, 92190 Meudon Principal Cedex, France

† Lunar and Planetary Laboratory, University of Arizona, Tucson, Arizona 85721, USA

‡ Université de Paris 7, France

§ NASA Johnson Space Center, Houston, Texas 77058, USA

|| European Southern Observatory, La Silla, Chile

The occultation of a star by Neptune on 20 August 1985 was observed at 2.2 μm wavelength with telescopes at the European Southern Observatory (ESO) and the Cerro Tololo Inter-American Observatory (CTIO). The detection of a 'central flash' midway between immersion and emersion has allowed the determination of Neptune's oblateness, ϵ , and the atmospheric extinction at 2.2 μm , which is related to the stratospheric methane mixing ratio. We find $\epsilon = (2.08^{+0.19}_{-0.18}) \times 10^{-2}$ and, assuming a stratospheric temperature of 120 K, infer a value of 0.6% (with an uncertainty of a factor of 10) for the methane mixing ratio (CH_4/H_2) at 0.3 mbar. The latter value may indicate supersaturation of methane in Neptune's stratosphere.

STELLAR occultations provide an opportunity to study upper planetary stratospheres and to determine planet oblatenesses; they also allow the detection of ring or arc systems around giant planets. Here we report observations of the occultation by Neptune of star no. 39 in the list by P. D. Nicholson, K. Matthews and G. Gilmore (personal communication) on 20 August 1985, made with the ESO 1-m telescope and the CTIO 1.5-m telescope. Both observations were made in the K band and with a sampling frequency of 100 Hz. An unexpected highlight of this experiment was the observation of a 'central flash' midway between immersion and emersion, when the star was behind the centre of Neptune. This event allows the independent determination of Neptune's oblateness and stratospheric methane mixing ratio. Here we focus only on the study of the flash. Isothermal fits and inversion of the immersion and emersion curves will be presented elsewhere, together with details of an arc structure detected around Neptune at Mauna Kea during the same night.

Observations

Figure 1 shows the occultation light curves and the structure of the observed flashes. A strong peak ($\sim 20\%$ of unocculted stellar flux at ESO and $\sim 15\%$ at CTIO) appears first, followed by a wide plateau with several bumps and then a double-peaked structure, after which the signal decreases. The structures observed at the two sites are well correlated (see also Fig. 5).

Theory and modelling of the flash

Theory. The central flash, produced when the star was behind the central part of Neptune, is due to curvature of the limb perpendicular to the light of sight, which causes convergence of the light to the central area of the geometric shadow of the planet¹. If, as a first approximation, we consider the planet to be spherical, we see that when the star is exactly behind the centre of the disk, the atmospheric refraction is radially symmetrical: the atmosphere is a spherical lens which focuses the light rays on its axis. For on-axis positions, the focusing of the rays is more efficient than spreading due to differential refraction, and this results in a sudden increase of the intensity (Fig. 2).

Because Neptune is significantly oblate, the limb of the shadow is approximately elliptical, and the atmosphere acts as an elliptical lens. The locus of perfect focusing is no longer restricted to the axis of the lens, but its projection in a plane parallel to the plane of the shadow is a curve called the evolute (or the caustic) of the ellipse. The light rays that fall near a given point on the limb converge to the centre of the curvature

of the ellipse at the point concerned, and the caustic is then formed by the centres of curvature of the ellipse (see Fig. 15 of ref. 1).

Letting r_e and r_p be the equatorial and polar radii of the planet at the level probed by the flash, and B ($= -25.04^\circ$) the planetocentric declination of the Earth, the equation of the limb of the shadow is:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad (1)$$

where a , the semi-major axis, is equal to r_e , and b , the semi-minor axis, is given by

$$b = \frac{r_e r_p}{\sqrt{r_p^2 \sin^2 B + r_e^2 + \cos^2 B}} \quad (2)$$

The apparent oblateness of the shadow is then

$$\epsilon' = \frac{a-b}{a} = 1 - (1-\epsilon)(\cos^2 B + (1-\epsilon)^2 \sin^2 B)^{-1/2} \quad (3)$$

where

$$\epsilon = \frac{r_e - r_p}{r_e} \quad (4)$$

is the true ellipticity of the planet.

Modelling. To calculate the theoretical profile of the flash, we assume the atmosphere to be isothermal in the region probed by the flash (in particular, any horizontal gradient is neglected) and we use the following method¹. For a point $M(x, y)$ within the shadow, the intensity $\phi(x, y)$ (normalized to the unocculted flux) is the sum of the contributions from the two to four M_i points of the limb whose normals intersect $M(x, y)$. The contribution from each point M_i is evaluated by solving the Baum and Code² equation with a scale height $H = RT/mg$ (where R is the gas constant, T is the temperature, m is the mean molar mass and g is the acceleration of gravity) and, to account for the curvature of the limb, the flux thus obtained is multiplied by the ratio of the radius of curvature of the ellipse at M_i to the distance $d = MC_i$, where C_i is the centre of curvature of the ellipse at M_i . As expected for caustics, the intensity becomes infinite on the evolute if one does not take into consideration the finite size of the star (in this case, 6.5 km radius at Neptune's distance, as estimated from the occultation by the Neptune arc observed from the Canada France Hawaii Telescope (B. Sicardy *et al.*, in preparation). To avoid tedious calculations, the profile of the flash is smoothed in one dimension only, along the most

sensitive direction—that is, perpendicular to the evolute. Finally, to allow for atmospheric absorption along the path of the rays, we multiply the whole flash profile by a transmission factor $e^{-\tau}$. The intensity throughout the central flash profile is approximately proportional to the scale height near the ray's closest approach to the planet; the effect on the profile of a transmission factor $e^{-\tau} < 1$ is therefore indistinguishable from the effect of a reduction of the scale height by the same factor. In the absence of multiwavelength data, it is necessary to use independent information about the scale height in the region probed by the flash to decide whether an inferred value of the transmission factor represents a finite τ or instead a lower temperature.

Atmospheric level probed by the flash. For an isothermal atmosphere above the probed level, the refractivity of the atmosphere at the level probed by the flash is given by³

$$\nu_t = \frac{\theta(h_t) H^{1/2}}{(2\pi r_e)^{1/2}} \quad (5)$$

where $\theta(h_t) = r_e/D$ is the light-ray bending angle at the altitude probed by the flash (h_t), and D is the Earth-planet distance (4.44×10^9 km). The number density at the altitude probed by the flash is given by

$$n_t = \frac{L\nu_t}{\nu_{\text{STP}}} = \frac{L}{\nu_{\text{STP}}} \frac{1}{D} \sqrt{\frac{Hr_e}{2\pi}} \quad (6)$$

where L is the Loschmidt number ($L = 2.687 \times 10^{19}$ molecules cm^{-3}), and ν_{STP} is the refractivity of the atmosphere at standard temperature and pressure. For $g = 1,090 \text{ cm s}^{-2}$, and an atmosphere of 90% H_2 and 10% He , we obtain $m = 2.2 \text{ g}$, so $H(\text{km}) = T(\text{K})/2.88$, and $\nu_{\text{STP}} = 1.26 \times 10^{-4}$ at $2.2 \mu\text{m}$ (ref. 3). The pressure at the probed level is then $p_t = n_t kT$, where k is the Boltzmann constant.

The column density N_t of the atmosphere traversed by the light rays is

$$N_t = n_t \sqrt{2\pi r_e H} \text{ molecules cm}^{-2} = \frac{Hr_e}{\nu_{\text{STP}} D} \text{ km atm} \quad (7)$$

with all the distances expressed in kilometres. Here $N_t = 1.564 \times 10^{-2} T \text{ km atm}$.

Fitting the flash profile

Oblateness of Neptune. Once the astrometry is fixed, the shape and height of the flash are determined by the four parameters r_e , ϵ , H and $e^{-\tau}$. Both the radius and the oblateness of the planet affect the profile of the flash by changing the size of the evolute (its typical dimension is $2\epsilon r_e$) and thus its position relative to the path of the telescope in the shadow. The flash profile is much more sensitive to ϵ than to r_e , and the uncertainty in ϵ ($\sim 20\%^{4,5}$) is much greater than that in r_e (0.2%). We therefore fix r_e at 25,200 km in the calculations.

Occultation measurements have provided temperatures of $\sim 150 \text{ K}$ at the $1 \mu\text{bar}$ level⁵⁻⁷, and far-infrared photometry indicates that the temperature is $\sim 60 \text{ K}$ at the 40 mbar level^{8,9}. Appleby's¹⁰ radiative convective models show that the temperature can vary from 70 to 150 K at the 0.1–1 mbar level, depending on the amount of aerosol heating and possible CH_4 supersaturation, but also that the temperature at this level must be $> 100 \text{ K}$ to account for the high brightness temperatures observed at 20 and $35 \mu\text{m}$. For our 'nominal' model, we adopted $T = 120 \text{ K}$; the corresponding pressure is then 0.32 mbar.

We started with prediction astrometry provided by P. D. Nicholson (personal communication), which was derived from published occultation mid-times and Neptune shape parameters^{4,5}. We used Neptune ephemeris data from the *Annuaire du Bureau des Longitudes* (1985) (P. Bretagnon, personal communication). To allow for the uncertainty in this initial astrometric solution, two additional parameters were left free: a possible shift Δs between the predicted and actual position

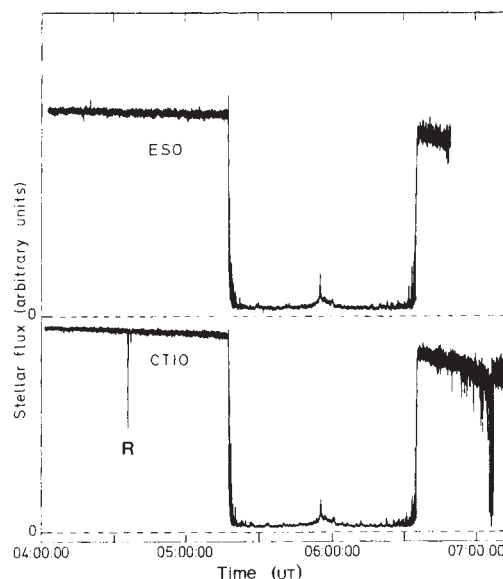


Fig. 1 Occultation light curves and the central flash as observed from ESO and CTIO. Each light curve is plotted at 1-s time resolution. The excursion labelled 'R' is caused by re-centering of the telescope.

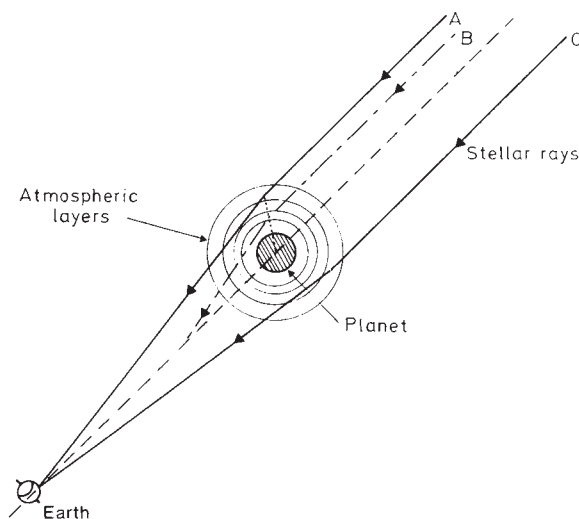


Fig. 2 Schematic explanation of the central flash for a spherical planet. Rays that traverse the atmosphere at different depths, such as A and B, are spread by differential refraction; in contrast, two rays which traverse the atmosphere at the same depth (such as A and C) converge to the same point on the axis and produce the flash.

of the observatory along the motion of the shadow, and the 'impact parameter' p_0 , which is the distance of closest approach to the centre of the shadow.

Our method for fitting the flash was as follows. For a reasonable set of parameters ($\epsilon = 0.021$; $T = 150 \text{ K}$; $e^{-\tau} = 1$; position angle of the projected north pole $P_n = 22.9^\circ$, as determined from the angular momentum of the Neptune-Triton system, assuming a Triton to Neptune mass ratio of 0.00128 (ref. 11)), the theoretical profile was calculated, and cross-correlated with the observed profiles, to obtain the central flash times. We obtained $t = 05:55:19 \text{ UT}$ for the ESO flash and $t = 05:55:21.0 \text{ UT}$ for the CTIO flash. Using these times, P_n was determined to be 20.1° . A simultaneous least-squares fit to the ESO and CTIO flash profiles with P_n as an additional parameter yielded $P_n = 20.0^\circ$, in good agreement with the results of the first method. As a

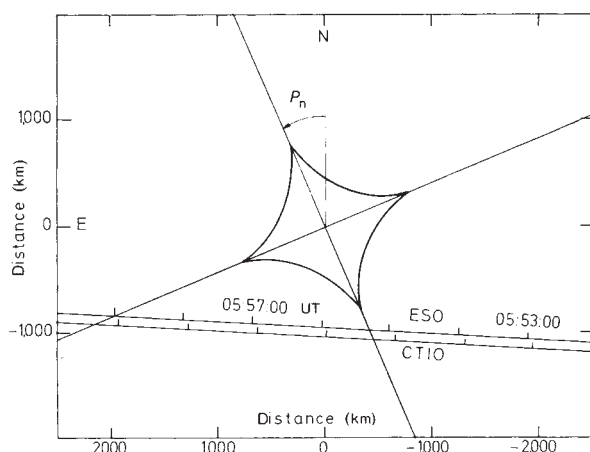


Fig. 3 Path of ESO and CTIO relative to Neptune's evolute as projected in the plane of the sky. For this diagram, r_e is taken to be 25,200 km, the apparent oblateness of the shadow, $\epsilon' = 0.0171$ and $P_n = 20.1^\circ$.

further check on the uncertainty in P_n , we measured the times of the peak central intensity at both stations by performing a least-squares fit of a gaussian profile to only those data points near the peak. This method yielded $P_n = 19.3^\circ$. We estimate the uncertainty in determination of central peak times to be 0.2 s, and the uncertainty in P_n to be 1° . Assuming a zero mass for Triton, Harris¹¹ ephemeris would yield $P_n = 21.6^\circ$. The disagreement with Harris¹¹ ephemeris thus appears to be significant, but must be taken with some caution as small atmospheric density fluctuations can cause significant systematic distortions of the central intensity profile. The shift Δs was determined to ensure that each observatory crossed the evolute axis at the time of the flash. Then, for a grid of points (ϵ, p_0), theoretical profiles were calculated for a transmission equal to unity, and for each point of the grid, a linear regression was made to find the transmission and the zero level that best matched the observed curve, and the corresponding residuals were calculated. The best fit was obtained for a particular set of parameters ϵ, p_0 and $e^{-\tau}$.

Figure 3 shows the paths of the observatories with respect to the evolute in the plane of the shadow; Table 1 summarizes the results for the two observatories. As explained above, the factor $e^{-\tau}$ for any assumed temperature (in K) can be scaled from the results presented in Table 1 by multiplying the tabulated values by $120/T$. Note that the agreement between the two observatories is satisfactory. Figure 4 shows the quality of the fit in the central part of the flash. Our determination of Neptune's oblateness at ~ 0.3 mbar is $\epsilon = 0.0208^{+0.0019}_{-0.0018}$. The location of the centre of the planet, as deduced solely from our model fit to the central flash, is in excellent agreement with the oblateness and centre location determined from the limb profile fitted solely to the edge occultation points (discrepancy in centres is $< H$; data to be published elsewhere).

As shown in Fig. 5, the model predicts a weak secondary peak at $\sim 05:58$ UT, caused by the approach of the occultation tracks to the eastern tip of the evolute pattern (Fig. 3). Although the data show a general plateau in this region, secondary peaks in the data do not appear at the predicted time. Adjusting the model so as to force a detailed fit only to the secondary peaks violates other constraints. Thus, for example, it becomes impossible to fit the main peak correctly, and the location of the planet's centre must be moved away from agreement with the limb profile centre. It is possible to argue (to be presented elsewhere), that the secondary peaks in the central flash region are caused by modest deviations of atmospheric refractivity from the smooth, symmetric model assumed here, and that these peaks are produced by essentially the same type of atmospheric disturbances

responsible for the well-known rapid intensity fluctuations seen during immersion and emersion. Thus a least-squares fit to a model representing the mean properties of the atmosphere should yield valid results, although it will not fit the data in every detail. In particular, because scattering by refractivity fluctuations must conserve energy, the mean intensity level in the entire flash region will be largely unaffected by the fluctuations, and the inferred value of $e^{-\tau}$ will be correspondingly conserved.

Methane mixing ratio in Neptune's stratosphere. The inferred opacity along the path followed by the light rays is related to Neptune's 2.2- μm absorbers. Note that Rayleigh scattering provides negligible extinction (10^{-4} – 10^{-5}) at 2.2 μm for paths of a few km atm (ref. 12), as is also the case for scattering by fine aerosols^{12,13}, and that at the pressure probed by the flash (< 0.5 mbar), H_2 – H_2 and H_2 –He pressure-induced absorptions are negligible. Thus, the observed absorption must be attributed entirely to methane.

In the region of transmission of the filter ($\lambda = 2.18 \mu\text{m}$, $\Delta\lambda = 0.42 \mu\text{m}$) the CH_4 rotation-vibration bands are those of the so-called octade. Three of the eight bands of the octade ($\nu_1 + \nu_4$, $\nu_3 + \nu_4$, $\nu_2 + \nu_3$) have been well studied, the others being much weaker^{14–16}. Only these three bands were considered here. For each of the three bands, each P, Q and R sub-branch shows J -multiplets. Integrated intensities of each multiplet up to $J \approx 14$ were provided by the Laboratoire de Spectrométrie Moléculaire at the Université de Dijon (J.-C. Hilico, personal communication), together with the number of lines involved in each multiplet. Because of the large path (~ 2 km atm) we had to consider even weaker lines (up to $J \approx 18$), whose intensities were evaluated by assuming they are governed by the exponential term, $\exp(-hcE/kT)$ (where h is Planck's constant, c is the speed of light and E is the lower energy level of the transition). Because of the quickly increasing number of lines in each J -multiplet, we used for each P, Q and R J -multiplet of each band, a band model of non-overlapping independent lines of unequal intensities, assuming an exponential distribution of the intensities¹⁷. For a given multiplet, where σ is the mean line intensity, the probability that a line has an intensity between S and $S + dS$ is $\sigma^{-1} e^{-S/\sigma} dS$, and the mean transmission over the band is $T = \exp(-\bar{A})$, where

$$\bar{A} = \frac{1}{\delta} \int_0^\infty e^{-S/\sigma} dS \int_{-\infty}^\infty [1 - \exp(-S f_\nu a)] d\nu \quad (8)$$

in which δ is the mean spacing of the lines, a is the column density of the absorber and f_ν is the line profile. In our case,

Table 1 Results of the isothermal fit

	ESO	CTIO
Time of flash (UT)	05:55:19.1	05:55:21.0
Sky-plane coordinates at time of flash*		
ζ_g (km)	–5,746.7	–5,726.5
η_g (km)	1,378.0	1,379.3
ζ_p (km)	5,395.3	5,344.0
η_p (km)	–2,341.0	–2,424.6
Shift $\Delta s^\dagger$ (km)	—	1
Shift $\Delta p_0^\ddagger$ (km)	—	–2.5
Oblateness	$0.0203^{+0.0011}_{-0.0013}$	$0.0212^{+0.0015}_{-0.0017}$
Transmission§	$0.72^{+0.18}_{-0.14}$	$0.67^{+0.21}_{-0.16}$
Methane mixing ratio§		
Upper limit	2×10^{-2}	4×10^{-2}
Lower limit	1×10^{-4}	2×10^{-4}
Most probable value	4×10^{-3}	8×10^{-3}

* Subscript g refers to geocentric coordinates; and subscript p refers to displacements due to parallax effects.

† With respect to ESO shift.

‡ Correction to impact parameter, with respect to ESO correction.

§ Assuming a stratospheric temperature $T = 120$ K.

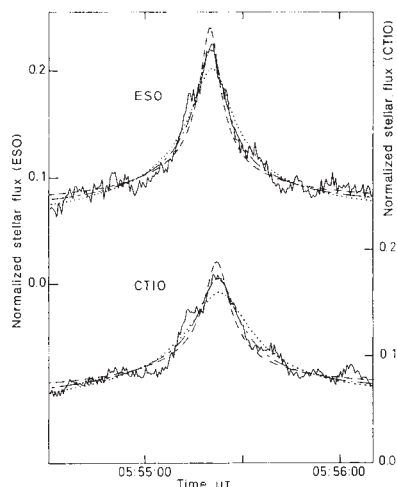


Fig. 4 Fitting of the central part of the flash for various values of ϵ . For ESO: —, $\epsilon = 0.0214$; ---, $\epsilon = 0.0203$; ···, $\epsilon = 0.0190$. For CTIO: —, $\epsilon = 0.0227$; ---, $\epsilon = 0.0212$; ···, $\epsilon = 0.0195$. The data (—) are plotted at a time resolution of 0.5 s.

the Doppler width $\alpha_D = 4.3 \times 10^{-7} \nu \sqrt{T/M}$ (where M is the molar mass of CH_4) is much greater than the Lorentz width α_L (typically $\alpha_D = 6 \times 10^{-3} \text{ cm}^{-1}$ versus $\alpha_L = 3 \times 10^{-3} \text{ cm}^{-1}$), and the cores of the lines have a Voigt profile¹⁸, the far wings being Lorentz-shaped.

For each multiplet, the integrated intensity and the number of the lines provides σ and δ . Finally, for each temperature and various methane mixing ratios, we obtained the general transmission over the filter as the product of the transmission of each multiplet.

The methane mixing ratio was adjusted to match the observed opacity, and our conclusions are summarized in Table 1. If $T = 120 \text{ K}$ (nominal model) we find a methane mixing ratio of 0.6% with an uncertainty of a factor of 10. If $T = 150 \text{ K}$, the methane mixing ratio is 1% with a factor of 4 uncertainty. If $T = 100 \text{ K}$, we can derive only an upper limit of 2%, with 0.1% as the most probable value.

Discussion

Oblateness. Measurements of planetary oblateness are of interest because they are related to internal structures through the gravitational fields. To first order, the following relation holds⁵:

$$\epsilon = \frac{1}{2} J_2 (3 + 1/\Lambda_2) \quad (9)$$

where J_2 is the gravitational quadrupole moment, and Λ_2 is the second-degree response coefficient as defined by Hubbard¹⁹: the smaller the value of Λ_2 , the more centrally condensed is the planet.

Recent determinations of ϵ were obtained from the results of the 15 June 1983 occultation by Neptune. From separate analyses of astrometric solutions for various stations, Hubbard *et al.*⁵ derived $\epsilon = 0.022 \pm 0.004$, whereas French *et al.*⁴ found 0.0191 ± 0.0017 . We find here an intermediate value of $\epsilon = 0.0208^{+0.0019}_{-0.0018}$, which is in fairly good agreement with these two determinations. It must be emphasized that, due to the very unusual path of the observatories in the shadow (they passed very close to the evolute but did not cross it), the shape of the flash is very sensitive to the oblateness of the model, which accounts for the small uncertainty in our determination of ϵ .

J_2 is equal to 0.0037–0.0043, as determined from the motion of Triton¹¹. Such values yield $\Lambda_2 = 0.12$ –0.15, which would imply that Neptune is less condensed than Uranus ($\Lambda_2 = 0.09$).

Methane mixing ratio. Since the discovery of emission features due to methane at $7 \mu\text{m}$ in the spectrum of Neptune²⁰, there

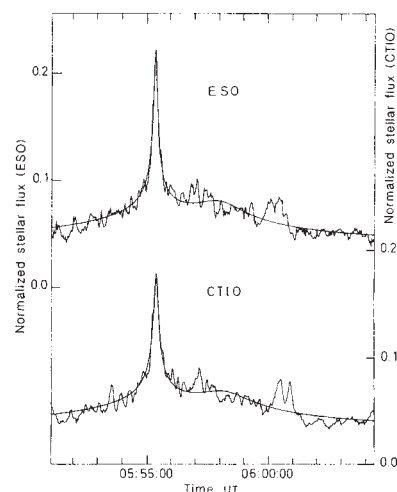


Fig. 5 Fitting of the entire flash profile. For ESO, $\epsilon = 0.0203$; for CTIO, $\epsilon = 0.0212$. The data are plotted at a time resolution of 2 s.

have been strong indications that methane could be supersaturated in Neptune's stratosphere. Even before this observation, Macy and Trafton²¹ had suggested methane supersaturation from brightness temperatures in the $20\text{-}\mu\text{m}$ region, and had found a stratospheric methane mixing ratio of 2×10^{-3} . Courtin *et al.*²² also derived a methane mixing ratio of 2×10^{-3} , from analysis of the $8\text{-}\mu\text{m}$ spectrum of Neptune. However, Gillett and Rieke²³ suggested that the high brightness temperatures at $8 \mu\text{m}$ can be explained in terms of a high stratospheric temperature instead of methane supersaturation, and they considered the possibility of a methane mixing ratio as low as 10^{-6} , with a stratospheric temperature of 150 K . Macy²⁴ reanalysed these works, and derived a methane mixing ratio of 10^{-2} – 10^{-3} in the stratosphere. More recently, Orton *et al.*²⁵ reported the first unambiguous detection of Neptune in the 10.3 – $11.4\text{-}\mu\text{m}$ region and derived a stratospheric methane mixing ratio of 0.01 – 0.1 , probably approximately equal to the tropospheric value. Our observation, which for an assumed stratospheric temperature of 120 K implies $\text{CH}_4/\text{H}_2 = 0.1$ – 1% , also suggests a large enrichment of stratospheric methane, compared to the case where methane is trapped at the tropopause (assuming a tropopause temperature of 53 K (ref. 9), the stratospheric mixing ratio would be 5×10^{-5}). These results require that methane is migrating through the tropopause cold trap, and several explanations have been proposed^{8,25,26}: (1) injection of methane into the stratosphere by strong, localized atmospheric convective motions; (2) transport of small methane particles into the warm stratosphere where they re-sublime, upward eddy diffusion overcoming the sinking of particles due to gravity; (3) an atmosphere so stable and homogeneous and nucleation sites so rare that methane remains supercooled in the inversion region; and (4) the existence of substantial areas where the inversion level is much warmer (leaks in the cold trap). Note that Appleby's²⁷ Neptune model suggests that the stratospheric methane abundance is near local saturation, a mechanically unstable situation that could, by temperature perturbations, produce an 'inverted cloud'.

Conclusion

We have here derived values for Neptune's oblateness and stratospheric methane mixing ratio from an isothermal model of the central flash profiles. We emphasize that the oblateness is derived solely from the flash profiles, and does not take into account the emersion and immersion times. A determination of

Neptune's oblateness from emersion and immersion times will be presented elsewhere, and comparison of the results will be a good test for the two methods. Nevertheless, the fact that the times and shapes of the two flashes observed independently are compatible with the relative positions of the observatories (see Table 1) gives us confidence in the model presented here. This model, however, does not fit the complex structure of secondary bumps. The effects of non-isothermal wave-like structures in the

atmosphere are under study and will be presented elsewhere.

E.L. thanks Jean-Claude Hilico for providing the intensities of the CH₄ bands; W.B.H. and F.V. acknowledge the support of NASA grant NSG-7045 and thank Brooke Gregory and Patricio Ugarte for essential contributions to the CTIO experiment; and we thank J. L. Elliot for constructive comments. F.V. is a visiting astronomer at CTIO, supported by NSF contract AST78-27879.

Received 3 July; accepted 29 September 1986.

1. Elliot, J. L. *et al. Astrophys. J.* **217**, 661-679 (1977).
2. Baum, W. A. & Code, A. D. *Astr. J.* **58**, 108-112 (1953).
3. Wasserman, L. H. & Veverka, J. *Icarus* **20**, 322-345 (1973).
4. French, R. G. *et al. Astr. J.* **90**, 2624-2638 (1985).
5. Hubbard, W. B. *et al. Astr. J.* **90**, 655-667 (1985).
6. Veverka, J., Wasserman, K. & Sagan, C. *Astrophys. J.* **189**, 569-575 (1974).
7. French, R. G., Elias, J. H., Mink, D. J. & Elliot, J. L. *Icarus* **55**, 332-336 (1983).
8. Tokunaga, A. T., Orton, G. S. & Caldwell, J. *Icarus* **53**, 141-146 (1983).
9. Moseley, H., Conrath, B. & Silverberg, R. F. *Astrophys. J.* **292**, L83-L86 (1985).
10. Appleby, J. F. *Icarus* **65**, 383-405 (1986).
11. Harris, A. W. *NASA Conf. Publ.* **2330**, 357-373 (1984).
12. Fink, U. & Larson, H. P. *Astrophys. J.* **233**, 1021-1040 (1979).
13. Podolak, M. & Giver, L. P. *Icarus* **37**, 361-367 (1979).

14. Brown, L. R. *J. molec. Spectrosc.* **96**, 94-105 (1982).
15. Hunt, R. H., Brown, L. R., Toth, R. A. & Brault, J. W. *J. molec. Spectrosc.* **86**, 159-169 (1981).
16. Hilico, J. C., Degni, J., Champion, J. P. & Guelachvili, G. *J. molec. Spectrosc.* **81**, 277-302 (1980).
17. Goody, R. M. *Atmospheric Radiation* (Oxford University Press, 1964).
18. Rao, K. N. *Molecular Spectroscopy: Modern Research* (Academic, New York, 1976).
19. Hubbard, W. B., *NASA Conf. Publ.* **2330**, 291-325 (1984).
20. Macy, W. & Sinton, W. *Astrophys. J.* **218**, L79-L81 (1977).
21. Macy, W. & Trafton, L. M. *Icarus* **26**, 428-436 (1975).
22. Courtin, R., Gautier, D. & Lacombe, A. *Icarus* **37**, 236-248 (1979).
23. Gillett, F. C. & Rieke, G. H. *Astrophys. J.* **218**, L141-L144 (1977).
24. Macy, W. *Icarus* **41**, 153-158 (1980).
25. Orton, G. S., Tokunaga, A. T. & Caldwell, J. *Icarus* **56**, 147-164 (1983).
26. Orton, G. S. & Appleby, J. F. *NASA Conf. Publ.* **2330**, 89-155 (1984).
27. Appleby, J. F. thesis, State Univ. New York at Stony Brook (1980).

LETTERS TO NATURE

Have burnt-out galaxies and galaxy clusters been detected?

Joseph Silk

Institut d'Astrophysique de Paris, 98 bis, Boulevard Arago,
F-75014-Paris, France and Department of Astronomy,
University of California, Berkeley, California 94720, USA

Close pairs of quasars are receiving much attention for two reasons. The image doubling may be due either to gravitational lensing of a single quasar by an intervening object or to physical clustering, that is, spatially close quasars with similar redshifts. Here I show that an elevated rate of star formation may be responsible for burning out protogalaxies in the densest galaxy clusters, leading both to formation of candidate lenses and to an enhanced probability of close association of quasars in regions of apparently low galaxy density.

Interpretation of identical multiple images of quasars as gravitational lenses requires the lensing objects to be at least as massive as a galaxy^{1,2} and in one recent case a galaxy cluster³, but in several cases to lack the optical luminosity one would ordinarily expect. For the most extreme example, the quasar pair³ 1146+111 B,C with a separation of 157 arc s, the hypothesized lens mass is $\geq 10^{15} M_{\odot}$. In the absence of any very rich galaxy cluster, candidate lenses include a cosmic string^{4,5} and a supermassive black hole⁶. The recent discovery⁷ of spectral differences between the image pair 1146+111 B,C however, suggests that one may in this case be seeing a physically-associated quasar pair. This is not easy to understand unless quasars are much more common in large groups or rich galaxy clusters at redshift $z \approx 1$ than at the present epoch⁸. An equally serious problem is the lack of candidate objects of galactic mass for lensing some of the several arc-second separation quasar pairs: alternatively, in these cases too, the physical association of quasars must have been enhanced.

A possible explanation arises naturally in standard galaxy formation models. I will argue that there is a fundamental bias against detecting luminous matter in the densest large-scale regions of the Universe, namely great clusters of galaxies. This provides a reasonably natural mechanism both for 'hiding' galaxies to provide candidate lensing objects and for enhancing the frequency of occurrence of the quasar phenomenon in such regions.

Consider a situation in which galaxy clustering developed hierarchically after galaxies formed, as in the cold dark matter model⁹, although this assumption is not crucial to my argument. Vigorous star formation at the present epoch is tidally induced, whether by generation of density waves excited by the tumbling or rotation of a central bar¹⁰ or by the close approach of a companion galaxy¹¹. Indeed, the most intense bursts of star formation that are identified with IRAS (the Infrared Astronomical Satellite) galaxies are associated with interacting galaxies¹². From the theoretical point of view, the non-linear response of a cold system greatly amplifies even weak tidal perturbations¹³. The net effect is to enhance gas cloud collisions and dissipation, initiate large-scale gravitational instabilities, and trigger star formation. Empirically, the IRAS studies suggest that the specific star formation rate is enhanced by one or even two orders of magnitude in the most intense star bursts, such as that in Arp 220, relative to the solar neighbourhood.

Little is known about star formation in protogalaxies. But it is virtually certain that when the old spheroidal stellar component formed, the specific star formation rate was anywhere between two and three orders of magnitude higher than it is today in the solar neighbourhood. It was necessary to have formed $\sim 10^{11} M_{\odot}$ of low-mass stars in a timescale that is estimated to be $\sim 10^9$ yr. It cannot be much less, because stellar evolution calculations and chemical evolution models require this time to build up the heavy elements that are seen in the old stars and to account for the spectral synthesis of elliptical galaxies¹⁴. It cannot be much longer, because the oldest stars in our Galaxy show no evidence of any spread in formation epoch¹⁵. It is tempting to identify the initiator of protogalactic star formation activity with the similar trigger that is observed today in starburst galaxies. Indeed, strong tidal interactions are inevitable if galaxies formed hierarchically from a Zel'dovich spectrum of primordial adiabatic, gaussian density fluctuations, in which small-scale power has not been suppressed, as in the cold dark matter model⁹. Moreover, all available evidence suggests that the protogalactic initial stellar mass function (IMF) spanned a similar mass range to the present day IMF^{16,17}, thereby helping to justify our confidence in taking over inferences about star formation in nearby galaxies.

One additional ingredient of star formation has received considerable attention. The efficiency of star formation measures the mass fraction of gas that has been converted into stars. In the solar neighbourhood, the long-lived, low mass ($\leq 1 M_{\odot}$) stars account for most ($\sim 85\%$) of the total star formation rate¹⁸. In regions of vigorous star formation, however, including the inner

galaxy¹⁹, the situation appears to be rather different. In particular, in starbursts, the ratio of the observed massive star formation rate to the total gas supply implies that relatively few low mass stars can be forming²⁰. The IMF cannot extend with any appreciable power below $2\text{--}3M_{\odot}$, and in some cases, $5\text{--}10M_{\odot}$. There are plausible, although hardly rigorous, theoretical arguments concerning why this should be so^{21,22}.

The following inference now becomes logical, and possibly inevitable. If there are any regions of the Universe so dense that tidal encounters between protogalaxies occurred at a rate large enough to stimulate star formation to a level up to and even beyond that seen in starbursts, then few very-low-mass stars could have formed before massive star formation had effectively depleted the net gas supply. If the star-forming phase in a proto-elliptical continues for $\sim 10^9$ yr, there is enough time to form many low-mass stars. But if it lasts only $\sim 10^8$ yr, low-mass stars may barely form at all. The evidence for this is largely circumstantial. Star formation bursts typically last $\sim 10^8$ yr, and fail to form low mass stars, whereas low mass stars do form over a much longer timescale, some several billions of years in the solar neighbourhood²³. This bimodality between short-lived massive and long-lived low-mass star formation provides an attractive explanation of the local paucity of metal poor G-dwarfs in the disk²⁴. Hence I conclude that in any region where the characteristic tidal interaction timescale between galaxies is as short as $\sim 10^8$ yr, one would be in grave danger of exhausting the gas supply without producing a significant number of low-mass stars. In 10^8 yr, most of the mass lost by stellar evolution of massive stars is recycled: if the low-mass cutoff in the IMF is above $\sim 10M_{\odot}$, essentially all the initial protogalaxy mass is cycled into compact remnants. With an IMF $dN(m)/dm \propto m^{-3}$ for massive stars²⁵, at most 20% of the initial mass is returned even if the low-mass cutoff is $\sim 5M_{\odot}$. The bulk of the initial mass is trapped in compact remnants. Evidently, with three or more generations of star formation, one would consume $\geq 90\%$ of the initial gas supply even if the star formation efficiency is $\sim 50\%$ (as is required to form bound star clusters). Any stars of lower mass would presumably not have had time to form in significant numbers because their formation rate is too low. Exactly why bimodal star formation occurs is not fully understood, but there is ample circumstantial evidence, both from our Galaxy and from other galaxies, that massive star formation occurs at a high specific rate without accompanying low-mass star formation, while low-mass star formation proceeds at a more leisurely pace, and is generally accompanied by some massive star formation^{19,24}. Alternatively, verification of a flat IMF as may be present in metal-rich galactic globular clusters²⁶ would show that some bound stellar systems are capable of eventually becoming seriously deficient in visible stars relative to those with a solar neighbourhood IMF²⁷. This directly suggests that metal-rich galaxies could burn out after successive episodes of massive star formation have exhausted the gas supply.

Consider the double quasar 1146+111 B,C with a separation of 157 arc s. I first assume that the multiple images are due to a non-singular gravitational lens. Placing the lens at the optimal redshift to minimize its inferred mass, together with the observed separation between images, specifies the gravitational potential well of the lensing object, apart from a geometrical uncertainty²⁸. Assuming the lens to be spherical, one infers a surface density $\Sigma \geq 1 \text{ g cm}^{-2}$ and virial velocity dispersion $V_c \approx 3,000 \text{ km s}^{-1}$. If it is indeed a cluster of galaxies, the galaxy crossing timescale is $V_c/G\Sigma \approx 10^8$ yr. This dynamical timescale is sufficiently short (a typical value for Abell clusters is $\sim 10^9$ yr) that one may expect tidal interactions to have had an important role in cluster evolution.

To evaluate the role of tidal heating, I use an effective cross-section for energy transfer derived from numerical simulations²⁹, $\sigma_E = 30r_e^2 (V_*/V_c)^3$, where r_e is the effective radius of a galaxy, taken to have a de Vaucouleurs profile and to all be identical,

and V_* is the isotropic r.m.s. stellar velocity dispersion within a galaxy. The heating rate is:

$$\dot{E}/E = 1.4(G\Sigma)^2 \sigma_E / r_e V_*^2 V_c = 0.7\Sigma^2 r_3 V_{*,300} V_{c,3000}^{-4} (10^9 \text{ yr})^{-1}$$

where $r_3 = r_e/3 \text{ kpc}$, $V_{*,300} = V_*/300 \text{ km s}^{-1}$ and $V_{c,3000} = V_c/3,000 \text{ km s}^{-1}$. The actual situation would probably have been even more extreme, because during the initial collapse of the cluster the galaxy velocity was $\ll V_c$ and protogalaxies had considerably larger effective radii. Evaluation of $\dot{E}/E (= n_c \sigma_E V_c)$ for parameters typical of the Coma clusters of galaxies ($n_c \approx 200 \text{ Mpc}^{-3}$ in the core³⁰) yields, for example, only $\dot{E}/E \approx 0.01(10^{10} \text{ yr})^{-1}$. Thus $\dot{E}/E$ is some three orders of magnitude larger for the hypothesized cluster that is lensing 1146+111 B,C.

In the case of cold dark matter with primordial density fluctuations described by the usual power spectrum $1\delta_k|^2 \propto k^n$, I find that for scales just going nonlinear at redshift z , $\dot{E}/E \propto (1+z)^{2(n+5)/(n+3)}$, or over rich cluster scales where $n \approx -1$ for a Zeldovich fluctuation spectrum, it follows that $\dot{E}/E \propto (1+z)^4$ provided $V_* < V_c$. The density of the Coma cluster core corresponds to turn-around at background density equal to about $(1/200)\rho_c \approx 10^{-27} \text{ g cm}^{-3}$, where ρ_c is the observed core density, or at $1+z \approx 6\Omega^{-1/3}$ if Hubble's constant $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$. Hence galaxy formation at $1+z \geq 10$ would have resulted in producing burnt-out galaxies only when clusters rather rarer (earlier forming) than Coma first collapsed. Abell clusters contain $\sim 5\%$ of luminous galaxies, and their frequency corresponds to their being $\sim 2\sigma$ fluctuations in a primordial gaussian fluctuation spectrum appropriate to cold dark matter models³¹. Clusters which would have induced burn-out are accordingly $\sim 3\sigma$ fluctuations, since $(\delta\rho)_c \approx (1+z_c)(\delta\rho/\rho)_{r.m.s.}$, and should be relatively rare, amounting at most to ~ 1 per every 5 Abell clusters over the same filtering scale (the number density of density peaks being proportional to^{32,33} $\nu^2 \exp(-\nu^2/2)$ for $\nu\sigma$ fluctuations if $\nu \gg 1$).

I conclude that galaxies in clusters as extreme as that inferred to be producing 1146+111B,C would have undergone extremely strong tidal interactions during the initial collapse of the cluster, as well as subsequently. Such clusters should be rare compared with Abell clusters, but still sufficiently frequent to lens quasars with reasonable probability. Galaxy burn-out may even be a common enough phenomenon to account for some of the lenses which produce quasars pairs with arc-second separations: a definite prediction of my model is that there should also be a rich cluster of burnt-out galaxies in such cases, so that broad separation ($\geq 100 \text{ arc s}$) doubling should be sought in the fields of putative lenses where no luminous galaxy candidate has hitherto been found.

Next, I consider the possibility that formation of burnt-out galaxies may also be associated with an enhanced frequency of occurrence of quasar pairs. The tidal interactions that are able to stimulate massive star formation are also capable of driving molecular clouds and stars onto radial orbits which fuel the massive black hole that may already have formed at the centre of a galaxy. Precisely when or how such massive holes form is a contentious issue, but it is generally believed that the fuelling of such a hole of $\sim 10^8 M_{\odot}$ at a rate $\geq 10 M_{\odot} \text{ yr}^{-1}$ is responsible for the quasar phenomena. Indirect evidence from absorption line studies suggest that quasars at $z \geq 1.2$ are indeed in rich clusters^{34,38}, while direct imaging reveals that low-redshift quasars apparently avoid rich clusters³⁹. As quasars evolve strongly in luminosity with redshift, this provides circumstantial evidence that the rich cluster environment may enhance quasar activity. The burnt-out galaxy hypothesis may render many of the associated galaxies sub-luminous; others, however, may be caught in their luminous protogalactic starburst phase. It is intriguing that the probable narrow emission line galaxy³⁵ at $z = 3.218$ that is a companion to the quasar PKS1614+051 is unusually luminous for a normal galaxy, and is perhaps the best candidate to date for a protogalaxy with the characteristic strong

$Ly-\alpha$ emission and compact appearance that models³⁶ predict.

Finally, I note that one prediction, if a rich galaxy cluster is responsible, either by locally enhancing the spatial frequency of quasars or by lensing, for the quasar pair 1146+111 B,C, is that gas infall and ensuing heating could lead, at least in the case of the lens hypothesis, to a potentially detectable Sunyaev-Zeldovich microwave background anisotropy³⁷ and, more seriously, to excessive hard X-ray emission; J. P. Ostriker and J.S. have estimated (unpublished) that the thermal bremsstrahlung X-ray flux should be $\sim 10^{11} (f_b/0.1)^2 \text{ erg cm}^{-2} \text{ s}^{-1} \text{ keV}^{-1}$ up to $\sim kT = 30 \text{ keV}$, where f_b is the mass fraction of the cluster in diffuse gas, and is probably well above the HEAO-1 all sky survey threshold if $f_b \approx 0.1$.

Burnt-out galaxies avoid this difficulty, because the available gas supply was mostly exhausted in forming massive stars. Any left-over gas would have participated in a cluster cooling flow, the high density of which should have led to strong cooling in the cluster core and galaxy formation in the same tidally stirred-up environment where the burnt-out galaxies previously formed. There should be some residual gas, however, at a density low enough that cooling should be unimportant (that is $f_b \approx 0.01$). This means that up to $\sim 10^{13} M_\odot$ of hot gas could be present, and it would contain processed matter, including dust. Thus the search for diffuse matter in the dark cluster that may be responsible for the quasar pair 1146+111 B,C should be pursued at X-ray, microwave and infrared wavelengths. More generally, the rarity of the burnt-out clusters guarantees that any contribution to the cosmic background light should be negligible. However, the possibility of observing vigorous starburst activity in the vicinity of high-redshift quasars deserves further attention. This activity could manifest itself, for example, in producing extended emitting or absorbing clouds of low ionization state and with velocity dispersions of up to $\sim 1,000 \text{ km s}^{-1}$ that are characteristic of vigorous starburst activity.

This research has been supported in part by NASA and by NSF. I thank Jean Audouze for his hospitality at I.A.P. where this work was completed.

Received 2 June; accepted 28 August 1986.

- Djorgovski, S. & Spinrad, H. *Astrophys. J.* **282**, L1-4 (1982).
- Weedman, D., Weymann, R. J., Green, R. F. & Heckman, T. M. *Astrophys. J.* **255**, L5-9 (1982).
- Turner, E. *et al.* *Nature* **321**, 142-144 (1986).
- Paczynski, B. *Nature* **319**, 567-568 (1986).
- Gott, J. R. *Nature* **321**, 420-421 (1986).
- Paczynski, B. *Nature* **321**, 419-420 (1986).
- Shaver, P. & Christiani, C. *Nature* **321**, 585-586 (1986).
- Bahcall, J. N., Bahcall, N. A. & Schneider, D. P. preprint (Institute for Advanced Study, 1986).
- Blumenthal, G., Faber, S., Primack, J. & Rees, M. J. *Nature* **311**, 517-525 (1984).
- Combes, F. & Gerin, M. *Astr. Astrophys.* **150**, 327-338 (1985).
- Icke, V. *Astr. Astrophys.* **144**, 115-123 (1985).
- Joseph, R. D. & Wright, G. S. *Mon. Not. R. astr. Soc.* **214**, 87-95 (1985).
- Goldreich, P. & Lynden-Bell, D. *Mon. Not. R. astr. Soc.* **130**, 125-158 (1965).
- Bruzual, G. *Astrophys. J.* **273**, 105-127 (1983).
- Sandage, A. *Astrophys. J.* **252**, 553-573 (1982).
- Silk, J. *Mon. Not. R. astr. Soc.* **205**, 705-718 (1983).
- Cayrel, R. *Astr. Astrophys.* (in the press).
- Miller, G. & Scalo, J. *Astrophys. J. Suppl.* **41**, 513-547 (1979).
- Gusten, R. & Mezger, P. *Vistas Astr.* **26**, 159-224 (1983).
- Rieke, G. H., Lebofsky, M. J., Thompson, R. I., Low, F. J. & Tokunaga, A. K. *Astrophys. J.* **238**, 24-40 (1980).
- Larson, R. *Mon. Not. R. astr. Soc.* **214**, 379-398 (1985).
- Silk, J. *Luminous Stars and Associations in Galaxies*, IAU Symp. **116**, 301-315 (Reidel, Dordrecht, 1986).
- Twarog, B. *Astrophys. J.* **242**, 242-259 (1980).
- Larson, R. *Mon. Not. R. astr. Soc.* **218**, 409-428 (1986).
- Scalo, J. *Fund. Cosm. Phys.* **11**, 1-278 (1986).
- McClure, R. D. *et al.* *Astrophys. J.* **307**, L49-53 (1986).
- Freeman, K. D. *The Evolution of Galaxies and Stellar Populations* (eds Tinsley, B. M. & Larson, R. B.) 133-148 (Yale University Observatory, 1977).
- Turner, E., Ostriker, J. P. & Gott, J. R. *Astrophys. J.* **284**, 1-22 (1984).
- Aguilar, L. & White, S. D. M. *Astrophys. J.* **295**, 374-387 (1985).
- Bahcall, N. *Astrophys. J.* **247**, 787-791 (1981).
- Kaiser, N. *Astrophys. J.* **273**, L17-L20 (1983).
- Doroshkevich, A. *Astrophys. J.* **304**, 15-62 (1986).
- Bardeen, J., Bond, J. R., Kaiser, N. & Szalay, A. *Astrophys. J.* **304**, 15-62 (1986).
- Weymann, R. J., Carswell, R. F. & Smith, M. G. A. *Rev. Astr. Astrophys.* **19**, 41-76 (1981).
- Djorgovski, S., Spinrad, H., McCarthy, P. & Strauss, M. A. *Astrophys. J.* **299**, L1-5 (1985).
- Meier, P. L. *Astrophys. J.* **207**, 343-350 (1976).
- Ostriker, J. P. & Vishniac, E. *Nature* **322**, 804 (1986).
- Foltz, C. B. *et al.* *Astrophys. J.* **307**, 504-534 (1986).
- Yee, H. C. K. & Green, R. F. *Astrophys. J.* **280**, 79-90 (1984).

A radio nebula associated with Circinus X-1

R. F. Haynes*, M. M. Komesaroff*, A. G. Little†‡, D. L. Jauncey*, J. L. Caswell*, D. K. Milne*, M. J. Kesteven*, K. J. Wellington* & R. A. Preston§

* Division of Radiophysics, CSIRO, PO Box 76, Epping, New South Wales 2121, Australia

† School of Physics Department, University of Sydney, Sydney, New South Wales 2006, Australia

§ Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109, USA

‡ Deceased.

The investigation reported here of the variable source Circinus X-1 (Cir X-1) reveals that it is embedded in a nebula of steady radio emission extending over several parsecs, orders of magnitude larger than the binary stellar system responsible for the fluctuating component of emission (detected at X-ray, radio and optical wavelengths). This is in marked contrast to most X-ray binaries, where an envelope of radio emission is conspicuously absent. There are difficulties in explaining the emission, but analogies with SS433 and the Crab nebula suggest possible models.

X-ray¹, optical², infrared^{3,4} and radio studies⁵ indicate that Cir X-1 is an eccentric double-star system, with one component of the binary being very compact, possibly a neutron star or a black hole. The 16.5-day periodicity, manifest across the observed frequency spectrum⁵, implies an orbital semimajor axis $\leq 0.5 \text{ AU}$. H I absorption measurements⁶ yield a firm lower limit to the distance of 8 kpc. At this distance the binary orbit subtends an angle $\leq 10^{-4}$ arc s, below the resolution limit of even very long baseline interferometry (VLBI). Nevertheless, Preston *et al.*⁷, using VBLI, reported a quiescent component of emission with angular size ≥ 0.2 arc s at 2.3 GHz and Haynes *et al.*² reported evidence of an extended radio source at 1.4 GHz.

The measurements presented here were made with the Molonglo Observatory Synthesis Telescope (MOST), which operates at 843 MHz; at the declination of Cir X-1 a 12-h synthesis observation⁸ yields a beam of 43×52 arc s. A contour diagram of Cir X-1 and the surrounding region is shown in Fig. 1.

The fact firmly established by the present results is that the variable compact source Cir X-1 is embedded in a radio nebula extending over ~ 5 arc min east-west and ~ 10 arc min north-south which shows no variability. The integrated flux of the nebula at 843 MHz is 0.56 Jy.

The compact object in Fig. 1 at RA 15 h 16 min 48.6 s and dec. $-56^\circ 59' 11.4''$ (1950 coordinates are used throughout this paper) coincides with the X-ray source¹. Table 1 compares our best estimate of its position with the position measured by Argue and Sullivan⁹ for the red star identified with the X-ray source and the position of the compact 2.3-GHz source measured using VLBI¹⁰.

The compact source is a well-known radio variable². Between May 1983 and June 1985 at 843 MHz we found that the flux

Table 1 Compact object associated with Circinus X-1

Object	Position (B1950)		Ref.
	RA	Dec.	
MOST compact object	15 h 16 min 48.58 $\pm$ 0.7 s	$-56^\circ 59' 11.4 \pm 0.5''$	This paper
Red star	15 h 16 min 48.436 $\pm$ 0.022 s	$-56^\circ 59' 11.4 \pm 0.28''$	9
2.3 GHz VLBI object	15 h 16 min 48.37 $\pm$ 0.04 s	$-56^\circ 59' 11.6 \pm 0.3''$	10

Uncertainties in the MOST positions are dominated by an assumed 0.5 arc s uncertainty in ionospheric refraction.

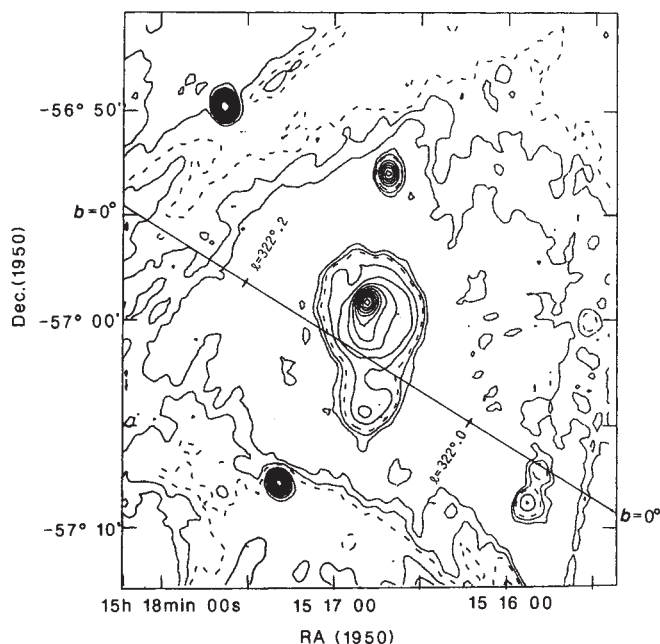


Fig. 1 843-MHz map of the Circinus X-1 region obtained by integrating seven 12-h synthesis maps made with the MOST between May 1983 and June 1985. Contours are shown at levels of 1, 3, 5 (dashed), 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85 and 90 mJy per beam area.

density varied between 40 and 88 mJy. The belief that the X-ray, radio and optical emission all emanate from Cir X-1 ultimately rests on the time variable correlation²; for example, detailed monitoring in April 1982 showed radio intensity changes clearly correlated with the 3–6-keV X-ray intensity over the same time (L. Kaluzienski, personal communication).

To study the detail of the nebula, we have subtracted the point radio source centred on Cir X-1 as shown in Fig. 2. The most prominent part of the nebula is approximately centred on the compact component and adjoins a weaker and narrower component to the south centred near RA 15 h 16 min 47 s, dec. $-57^{\circ}04'$. Earlier, lower-resolution measurements^{2,5} showed a quiescent source which we now assume arose predominantly from the diffuse nebula rather than from the compact variable source. These measurements may be combined with the present 843-MHz results ($S_{843} = 0.56$ Jy) to yield a distinctly non-thermal spectral index of $\alpha \approx -0.5$ ($S \propto \nu^{-\alpha}$). Non-thermal sources of this intensity have a spatial density of ~ 1 per 3 square degree, and Cir X-1 coincides with the nebula's centre to within 1 arc s; the probability of a chance coincidence is then ≤ 1 in 3,000. Thus, we conclude that the whole nebula is physically associated with Cir X-1 rather than a mere line-of-sight coincidence.

Cir X-1 has several characteristics in common with Cyg X-1 (refs 11, 12). However, as far as we know, Cyg X-1 displays no extended radio component comparable with that associated with Cir X-1.

In attempting to explain the extended distribution we consider two general types of model: (1) the nebula is the remnant of an earlier active phase which no longer persists. (2) The nebula has its energy supply maintained by continuous injection from a central engine. The first interpretation corresponds to the typical supernova explosion. In this case, we should expect to find a limb-brightened source. There is no evidence of this in Fig. 1, even when allowance is made for the limited angular resolution. Furthermore, the nebula has a surface brightness that would be anomalously low for a typical shell supernova remnant of this size¹³.

In the second kind of model, (2) the Crab nebula represents one variety of continuous injection model. In this example, the

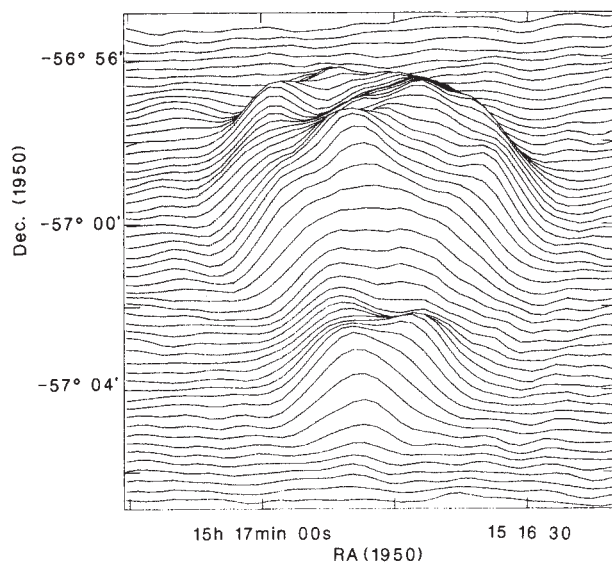


Fig. 2 Perspective view of the nebular region surrounding Cir X-1 obtained by subtracting the 843-Hz emission from the point radio source associated with Cir X-1 from the data shown in Fig. 1.

central engine is a rotating neutron star, and energy transfer to the surrounding nebula may be accomplished by low-frequency electromagnetic waves¹⁴ or by charged particles accelerated to ultra-relativistic energies in the enormous electric fields produced by the induction¹⁵. Either process requires that the central engine be in a highly evacuated cavity. Such processes are thus expected to be inoperative in Cir X-1 because it is an eccentric binary system^{16,17} with a high rate of accretion from a 'normal' star to a 'compact' star and the resulting mass flow may smother the energy transfer process. On the other hand, SS433 (refs 18, 19) another variable binary-star system, produces collimated jets which do successfully escape from the binary system and are not smothered by the mass transfer.

The Cir X-1 flares have the characteristics of the van der Laan expanding radio source model¹⁶. This leads us to suggest, as another variant of the continuous injection model, that the extended nebular is the accumulation of the energetic particles and plasma ejected during the flaring activity.

Clark *et al.* have argued that Cir X-1 is a runaway binary system. According to this theory Cir X-1 was ejected from the supernova G321.9–0.3, the centre of which is some 25 arc min south, at the time of its eruption. If this could be substantiated, it would rule out the possibility that the nebula surrounding Cir X-1 is itself a supernova remnant, and render a continuous injection model almost inescapable. The southward extension of the radio nebula would also fit naturally into such a picture. This model requires an asymmetric supernova explosion with a high ejection velocity. Accepting 8 kpc as a minimum distance to Cir X-1 and 10^5 yr as the age¹² of G321.9–0.3, we see that the transverse component of the ejection velocity needs to be 500 km s^{-1} . This is greater than the largest known pulsar velocity²⁰ (365 km s^{-1}).

There are several future lines of investigation which could resolve some of the questions raised by Cir X-1. First, proper-motion studies in the radio domain (using VLBI techniques) and perhaps in the optical region could determine whether Cir X-1 was ejected from G321.9–0.3. If so the nebula around Cir X-1 could not itself be a supernova remnant. We propose to undertake these studies using the Parkes-Tidbinbilla interferometer system. Second, further mapping of the nebula could be undertaken to investigate more fully the absence of limb-brightening and the spectral index. Finally, detailed theoretical modelling of the dynamics could yield an upper limit on the

time for circularization of the eccentric orbit and thus yield an upper limit on the lifetime of the system.

The MOST is operated by the University of Sydney under funding provided in part by the Australian Research Grants Committee, the University of Sydney Research Committee and the Science Foundation within the University of Sydney. We thank Professor B. Y. Mills and Dr J. M. Durdin for help with using the MOST. The research described in this paper was carried out in part by the Jet Propulsion Laboratory of the California Institute of Technology under contract to NASA.

Received 21 July; accepted 10 September 1986.

1. Kaluzienski, L. J., Holt, S. S., Boldt, E. A. & Serlemitsos, P. J. *Astrophys. J.* **208**, L71-L75 (1976).
2. Haynes, R. F. et al. *Mon. Not. R. astr. Soc.* **185**, 661-671 (1978).
3. Glass, I. S. *IAU Circ. No.* 2974 (1976).
4. Glass, I. S. *Mon. Not. R. astr. Soc.* **183**, 335-340 (1978).
5. Whelan, J. A. J. et al. *Mon. Not. R. astr. Soc.* **181**, 259-271 (1977).
6. Goss, W. M. & Mebold, U. *Mon. Not. R. astr. Soc.* **181**, 255-258 (1977).
7. Preston, R. A. et al. *Astrophys. J.* **268**, L23-L26 (1983).
8. Mills, B. Y. *Proc. astr. Soc. Aust.* **4**, 156-159 (1981).
9. Argue, A. N. & Sullivan, C. *Observatory* **102**, 4-6 (1982).
10. Argue, A. N., Jauncey, D. L., Morabito, D. D. & Preston, R. A. *Mon. Not. R. astr. Soc.* **209**, 11P-14P (1984).
11. Braes, L. E. & Miley, G. K. *Nature* **232**, 246 (1971).
12. Clark, D. H., Parkinson, J. H. & Caswell, J. L. *Nature* **254**, 674-676 (1975).
13. Milne, D. K. *Aust. J. Phys.* **32**, 83-92 (1979).
14. Ostriker, J. P. & Gunn, J. E. *Astrophys. J.* **157**, 1395-1417 (1969).
15. Goldreich, P. & Julian, W. H. *Astrophys. J.* **157**, 869-880 (1969).
16. Haynes, R. F., Lerche, I. & Murdin, P. G. *Astr. Astrophys.* **87**, 299-302 (1980).
17. Murdin, P. G. et al. *Astr. Astrophys.* **87**, 292-298 (1980).
18. Clark, D. H. & Murdin, P. G. *Nature* **276**, 45 (1978).
19. Margon, B. A. *Rev. Astr. Astrophys.* **22**, 507-536 (1984).
20. Anderson, B. & Lyne, A. G. *Nature* **303**, 597-599 (1983).

The historical record for Sirius: evidence for a white-dwarf thermonuclear runaway?

Frederick C. Bruhweiler*, Yoji Kondo† & Edward M. Sion‡

* Department of Physics, Catholic University of America, Washington, D.C. 20046, USA

† Laboratory for Astronomy and Solar Physics, NASA/Goddard Space Flight Center, Greenbelt, Maryland 20771, USA

‡ Department of Astronomy and Astrophysics, Villanova University, Villanova, Pennsylvania 19085, USA

Schlosser and Bergmann¹ recently presented evidence that in medieval times Sirius was a bright red star, rather than the present bluish-white star, from which they have suggested that Sirius B is a recently born white dwarf. However, their model poses severe evolutionary problems. We present the results of our attempts to detect possible planetary nebula ejecta toward Sirius, using data obtained by the International Ultraviolet Explorer (IUE) satellite. Based upon these results and in the light of recent advances in understanding white-dwarf evolution, we propose that Sirius B underwent a recent thermonuclear runaway event, triggered by a diffusion-induced CN reaction, to explain the historical behaviour of this star.

Two evolutionary problems remain in the Schlosser and Bergmann evolutionary picture. First, the implied cooling time for the white dwarf, Sirius B, is several orders of magnitude too short to be explained by current theories of white-dwarf cooling. For example, current predictions indicate that the time required for a white dwarf with an effective temperature $\sim 2.7 \times 10^4$ K, like Sirius B (ref. 2), to cool from 10^5 K, typical of a planetary nebula nucleus, is $\sim 10^7$ years. To explain such discrepancies would require either new physics or a rethinking of our current picture of the later stages of stellar evolution. Secondly, there

is no evidence of any stellar ejecta, although such ejecta could have escaped detection. A search for such ejecta is warranted, especially in view of recent analysis of ultraviolet spectral data obtained of the red-dwarf-white-dwarf binary system V471 Tauri³, which showed possible gaseous material moving away from that system at -590 km s^{-1} .

It is well known that interstellar ultraviolet absorption lines of cosmically abundant atomic species can provide a very sensitive means of detecting intervening gas. Thus, in our search for the putative ejecta, we have made use of data previously acquired by the IUE. These data come from the high-dispersion spectral image SWP 10075 obtained for the white dwarf Sirius B. These data cover the spectral region from $\sim 1,200$ to $2,000 \text{ \AA}$ and provide information on many ground-state transitions for several cosmically abundant species. A full discussion of the acquired data and a discussion of the observed interstellar features can be found elsewhere⁴.

The new observational results that we present here represent a reanalysis of the IUE image SWP 10075. This reanalysis is necessary because originally we were not looking for possible high-velocity interstellar absorption, and we can now apply additional techniques to place stronger detection limits on intervening gaseous absorption by using the almost featureless nature of white-dwarf spectra.

Our data analysis follows very closely that recently presented³. We have taken wavelength regions containing the strongest transitions for the abundant atomic species and converted these data to the velocity frame by using the transformation $v = c(\lambda - \lambda_0)/\lambda_0$, because the internal wavelength (and velocity) scale of the IUE cameras is extremely accurate ($\pm 2 \text{ km s}^{-1}$), absorption due to more than one ion at a particular velocity should be very closely reproduced at any velocity. The interstellar features of C II, Si II and O I are seen near zero velocity, but there is no evidence of other features within $3,000 \text{ km s}^{-1}$ of the rest velocity. There was a hint of a feature at the 1.5σ level near -30 km s^{-1} with respect to the interstellar component in the Si II data, but nothing is seen in other ions. Although planetary nebulae typically show expansion velocities $\sim 30 \text{ km s}^{-1}$, the recent discovery of sharp high-velocity features in V471 Tauri justifies searching for components over a large velocity interval. Because the species C II, Si II and O I have similar ionizations, we coadded these data in an attempt to enhance the detection limit in case weak high-velocity components of these ions existed. We found no evidence of additional absorption other than the expected interstellar features. The results of this coaddition are shown in Fig. 1. The same test was also performed upon the stronger members of the resonance doublets for the highly ionized species N V, C IV and Si IV. Again, no evidence of absorption was found either corresponding to the already observed low-ionization interstellar features or at higher velocities.

Based on these results, we find it hard to support the supposition that the Sirius stellar system has recently ejected any type of detectable nebulosity or gaseous shell. One might still argue that the velocity resolution (22 km s^{-1}) would not be sufficient if the expanding shell were 'masked' or unresolved from the intervening interstellar features. However, the implied average particle densities toward Sirius are in accord with those found in other short lines of sight in the local interstellar medium^{4,5}. We derive from the IUE data an upper limit of $3 \times 10^{17} \text{ cm}^{-2}$ for the total gas column density for the putative ejecta about Sirius B. If we assume an age of 1,000 years and an expansion velocity of 30 km s^{-1} , we get a corresponding upper limit of $3.6 \times 10^{-5} M_\odot$ for the mass of any ejecta. In addition, Copernicus results⁶, with about a twofold better resolution and much better signal-to-noise ratio, do not indicate multiple interstellar Mg II features at 2,795 and 2,802 Å. Thus, we must conclude that the physical mechanism causing the colour change of Sirius was due to a relatively quiescent process that did not involve any significant mass ejection.

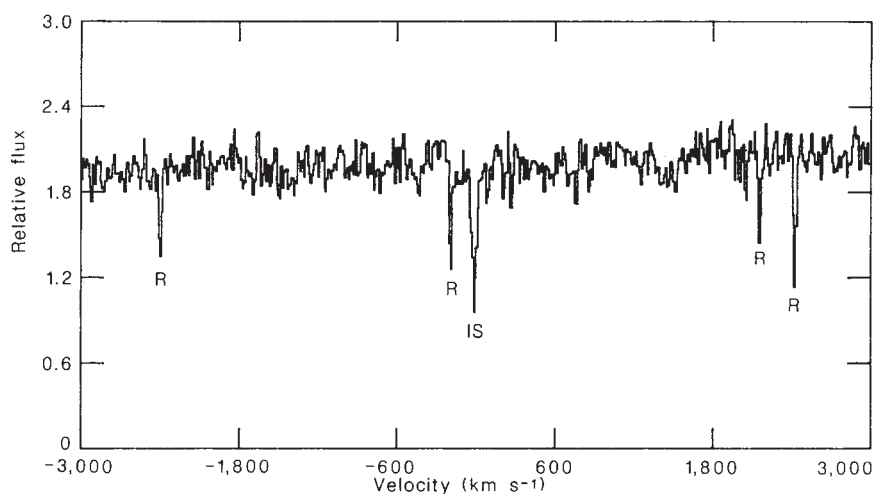


Fig. 1 Coadded data for regions corresponding to strong ground-state transitions of O I, C II and Si II. The coadded data for the O I 1,302.169, C II 1,334.532 and Si II 1,260.421 Å transitions are shown in the velocity frame. No evidence of absorption other than the expected interstellar contribution near zero velocity (labelled IS) is found. The other sharp features (R) are detector fiducial or réseau marks contributing from each of the data sets covering the respective atomic transitions. The data are from IUE image SWP 10075 acquired using the small aperture.

One other possible physical interpretation is that the red colour of Sirius during mediaeval times was due to intervening dust along the line of sight. Even though Sirius does not seem to have an infrared excess, at least two nearby A stars, Vega and Beta Pictoris, show strong evidence of possessing extended dust envelopes or disks⁷⁻¹⁰. The sizes of the dust particles in these disks are inferred to be large, possibly as large as 1 mm. These particles can provide extinction, but do not seem to yield any reddening at optical and ultraviolet wavelengths. Regardless of the properties of the dust, there cannot be reddening without extinction. To obtain an observed $B-V \geq 1.0$ as is deduced for Sirius in the sixth century would require substantial dimming. If we were to assume normal interstellar dust, Sirius would have been three magnitudes fainter than it appears today: it seems unlikely that Sirius would have attracted any particular attention as a bright star. In contrast, Schlosser and Bergmann find Babylonian sources that mention that Sirius was visible in the day time and must have been brighter than it is today. We must therefore find a more attractive explanation for the historical observations.

In recent years our picture of white-dwarf evolution has changed quite dramatically. There is now evidence that residual hydrogen burning is still occurring in hot-hydrogen-atmosphere (DA) white dwarfs like Sirius B. Recent theoretical efforts have demonstrated that nuclear shell burning, either by proton-proton reactions or by the diffusion-induced $^{12}\text{C}(p, \gamma)^{13}\text{N}$ can arise¹¹⁻¹³.

Burning by the $^{12}\text{C}(p, \gamma)^{13}\text{N}$ reaction is expected first to arise in DA white dwarfs like Sirius B ($T_{\text{eff}} = 27,000$ K, where T_{eff} is the stellar effective temperature) once they have cooled to temperatures less than about 30,000 K (refs 11, 12). This hydrogen-burning occurs as the diffusion tail of the hydrogen layer extends downward into the underlying helium layer, where it encounters the diffusion tail of carbon extending upward from the $^{12}\text{C}-^{16}\text{O}$ core.

The great sensitivity of the CN reactions to temperature could lead to a thermonuclear shell flash. However, the thermal stability of diffusion-induced reactions in a $1.05 M_{\odot}$ white dwarf like Sirius B has not yet been explored either with evolutionary-model calculations or by a linearized analysis of thermal stability. Here we extend the results of Iben and Tutukov¹³ and Michaud and Fontaine¹², who discuss white dwarfs of $0.6 M_{\odot}$ with $^{12}\text{C}-^{16}\text{O}$ cores, to the more massive Sirius B.

The model we present is that Sirius B did become a white dwarf some 10^7 years ago, as is generally accepted. However, it has cooled enough ($T_{\text{eff}} < 30,000$ K) to initiate a thermonuclear runaway via the $^{12}\text{C}(p, \gamma)^{13}\text{N}$ reaction. It has been shown^{14,15} that along its constant-radius cooling track where $\partial L / \partial M < 0$, a steadily burning white dwarf would be thermally unstable. In fact the steady-burning solutions show a high-luminosity stable branch and a low-luminosity unstable branch separated at the

point of maximum T_{eff} (bifurcation point) in the Hertzsprung-Russell diagram or in the $M_{\text{e}}-T_{\text{eff}}$ plane, where M_{e} is the mass of hydrogen above the burning region (see Figs 2 and 3 in ref. 16).

It is well known that the long-term behaviour of evolutionary models of actual white dwarfs, even with the steady-burning assumption removed, is still morphologically similar to that of the solutions of the steady-burning model shown in Fig. 3 of Iben^{16,17}. In other words, the evolution in the $M_{\text{e}}-T_{\text{eff}}$ plane as described in that figure is valid not only for accretion-induced thermal runaways, but also for a non-accreting star (or one with negligible hydrogen accretion compared with hydrogen lost by burning) experiencing a thermal runaway. The main effect of the hydrogen-burning runaway is to provide the energy necessary to allow the 'thin' hydrogen envelope to make a phase transition across a stability barrier from an unstable to a stable branch with very little change in the mass of the envelope.

Based on the above, we can describe the evolution of Sirius B during the proposed runaway event by applying the solutions or 'tracks' given by Figs 2 and 3 in Iben¹⁶ for a $1.1 M_{\odot}$ white dwarf. Therefore, if Sirius B did indeed undergo a thermal pulse, it then expanded to a giant structure, arriving on the stable branch, at a point determined by the mass of hydrogen lying above the burning region (M_{e}). If, as we assume, $M_{\text{e}} \approx 10^{-4} M_{\odot}$, Sirius B would have spent ~ 250 years mimicking a red giant, burning its envelope hydrogen stably along the plateau (stable) branch to the bifurcation point near maximum T_{eff} . From that point, it would proceed to cool rapidly, on a thermal timescale, downward to lower luminosity along the unstable branch. For larger M_{e} , longer times are spent on the stable branch. The expansion of Sirius B to a giant structure may have been accompanied by some mass loss but the higher thermal conductivity of a massive white dwarf such as Sirius B could have weakened the flash by the efficient conduction of energy below the burning region. Indeed, the upper limit ($3.6 \times 10^{-5} M_{\odot}$) for the mass of the ejecta about Sirius B tends to support this interpretation.

The qualitative model presented here is speculative, but considering the growing evidence of the importance of nuclear burning in white dwarfs, plus the physical and observational constraints placed upon the white-dwarf cooling times, models of thermal runaways in white dwarfs offer the best means of describing the historical data for Sirius B. We cannot dispose of the possibility of a thermal runaway triggered by accretion, but the current environment of Sirius B implies low accretion rates both from Sirius A and the interstellar medium. In any event, the historical records may offer a unique observational opportunity to test current theoretical models for white-dwarf evolution.

Note added in proof: An article by H. Takeuchi *et al.* in *Newton* 4, 138 (1984), mentions the Babylonian record of 700 BC, in which Sirius was observed as a red star. They consider the possibility that Sirius B was a red giant in 700 BC

but dismiss the idea because a red giant would not have evolved fast enough to become a white dwarf in only 2000 years. They speculate that Sirius B was a red giant tens of thousand years ago.

Received 27 May; accepted 26 August 1986.

- Schlosser, W. & Bergmann, W. *Nature* **318**, 45–46 (1985).
- Böhm-Vitense, E., Dettmann, T. & Kapranidis, S. *Astrophys. J.* **232**, L189–L192 (1979).
- Bruhweiler, F. & Sion, E. *Astrophys. J.* **304**, L74 (1986).
- Bruhweiler, F. & Kondo, Y. *Astrophys. J.* **259**, 232–243 (1982).
- McClintock, W., Henry, R., Linsky, J. L. & Moos, H. W. *Astrophys. J.* **225**, 465–481 (1978).
- Kondo, Y., Talent, D., Barker, E., Dufour, R. & Modisette, J. *Astrophys. J.* **220**, L97–L102 (1977).
- Aumann, H. H. *Bull. Am. astr. Soc.* **16**, 483 (1984).
- Aumann, H. H. *et al. Astrophys. J.* **278**, L23–L27 (1984).
- Smith, B. & Terile, R. *Science* **226**, 1421–1424 (1984).
- Kondo, Y. & Bruhweiler, F. *Astrophys. J.* **291**, L1–L5 (1985).
- Michaud, G., Fontaine, G. & Charland, Y. *Astrophys. J.* **280**, 247–254 (1984).
- Michaud, G. & Fontaine, G. *Astrophys. J.* **283**, 787–793 (1984).
- Iben, I. & Tutukov, M. *Astrophys. J.* **282**, 615–630 (1984).
- Sienkiewicz, R. *Astr. Astrophys.* **85**, 295–301 (1980).
- Papaloizou, J., Pringle, J. & MacDonald, J. *Mon. Not. R. astr. Soc.* **198**, 215–220 (1982).
- Iben, I. *Astrophys. J.* **259**, 244–266 (1982).
- Sion, E. & Starrfield, S. *Astrophys. J.* **303**, 130–135 (1986).

Ice core record of the $^{13}\text{C}/^{12}\text{C}$ ratio of atmospheric CO_2 in the past two centuries

H. Friedli, H. Löttscher, H. Oeschger, U. Siegenthaler & B. Stauffer

Physics Institute, University of Bern, Sidlerstrasse 5, CH-3012 Bern, Switzerland

The release of carbon into the atmosphere due to the activities of humans has caused an increase in concentration as well as a change in the isotopic composition of atmospheric carbon dioxide. CO_2 derived from fossil fuel combustion and from biomass destruction have $\delta^{13}\text{C}$ values of $\sim -25\%$ (compared to the atmospheric value of $\sim -7\%$) and are thus depleted in ^{13}C . We have measured $\delta^{13}\text{C}$ of CO_2 separated from air trapped in bubbles in ice samples from an ice core taken at Siple Station in Antarctica, in which it has been possible to demonstrate the atmospheric increase of CO_2 (ref. 1) and methane² with high time resolution. The isotopic results, together with the CO_2 record from the same ice core, yield information on the sources of excess carbon dioxide and provide a data base for testing the consistency of global carbon cycle models.

Prerequisites for obtaining a good ice core record of the recent past are a high accumulation rate which yields good time resolution, and the absence of summer melting (meltwater interferes with CO_2). Both conditions are fulfilled at Siple Station ($75^\circ 55' \text{ S}$; $83^\circ 55' \text{ W}$) where the mean annual air temperature is -24°C and the mean annual accumulation rate is 500 kg m^{-2} . The ice core which we studied was cut during the Antarctic summer of 1983–84 to a depth of 200 m by the Polar Ice Coring Office (Nebraska) and the Physics Institute of the University of Bern. Enclosure and age of the air have been discussed elsewhere^{1,3}. It has been possible to date the ice with an accuracy of ± 2 years down to a depth of 144 m (which corresponds to the year 1834) by counting seasonal variations in electrical conductivity⁴. Below that depth, the core has been dated by extrapolation.

The enclosed air is younger than the surrounding ice, because the enclosure of air in bubbles occurs only between depths of 64–76 m. At shallow depths atmospheric air still circulates through the open pores. In agreement with earlier papers^{1,2}, we adopted a value of 83 years for the age difference, and 22 yr for the width of the age distribution (time interval from 10–90% close-off of the total bubble volume).

The experimental procedure has been described previously^{5,6}. The air was extracted by grinding the ice samples (each $\sim 700 \text{ g}$)

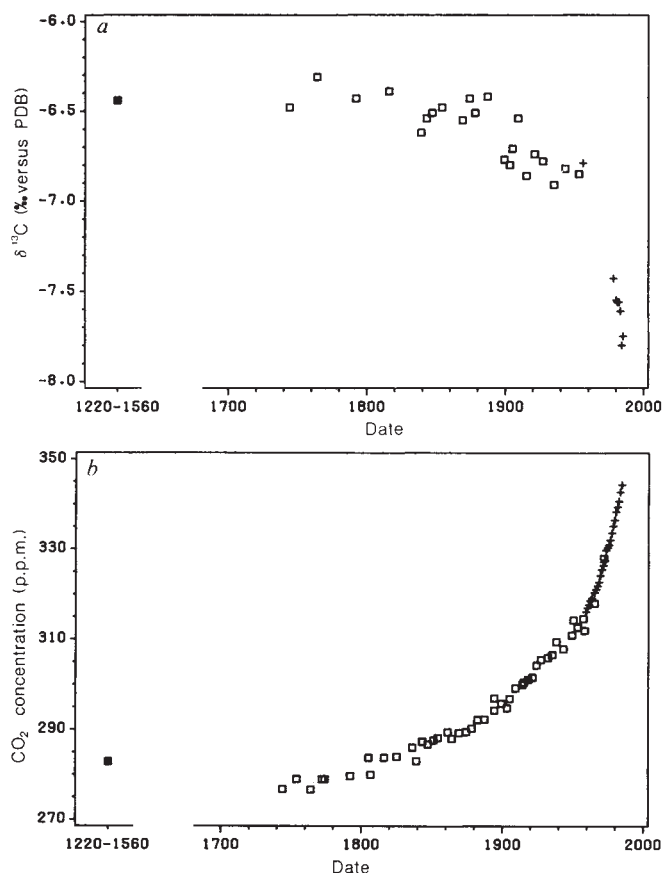


Fig. 1 a, $\delta^{13}\text{C}$ of CO_2 (in $\%$ versus PDB standard¹⁸ N_2O -corrected), b, CO_2 concentration in air extracted from ice cores from Siple Station (open squares) (data from this work and ref. 1) and from South Pole Station (black square)⁶. Also shown (crosses) are results from direct atmospheric samples at Mauna Loa (refs 8–10 and C. D. Keeling, personal communication).

at -20°C in an evacuated container with a milling cutter. The air escaping from the mechanically opened bubbles ($\sim 80 \text{ ml}$ STP per sample) was collected by condensation at 14 K in a small steel cylinder. Extraction efficiency is 80–90%, because not all bubbles are opened. A small amount of the evolved gas was used for gas-chromatographic measurement of the CO_2 concentration (accuracy $\pm 3 \text{ p.p.m.}$). CO_2 was quantitatively separated from the remainder of the sample by freezing it out at liquid nitrogen temperature in a metal vacuum line. Isotope ratios were measured by mass spectrometry, and the results were corrected for the presence of N_2O which has the same isotopic masses (44, 45 and 46) as CO_2 , following the procedure described by Mook and van der Hoek⁷, which we checked and confirmed by analysing artificial $\text{N}_2\text{O}/\text{CO}_2$ mixtures. For $\delta^{13}\text{C}$ corrections, the N_2O concentration in the CO_2 sample was determined by mass-spectrometry from the abundance of mass 30, which corresponds to the fragment NO^+ produced in the ion source. Details of this procedure will be reported in a separate paper. We observed variable and significantly higher N_2O concentrations than expected. Tests run afterwards showed that N_2O was produced in the vacuum line in a cold cathode manometer. Accuracy of the $\delta^{13}\text{C}$ results is 0.10%, as determined by analysing artificial CO_2 -in-air-mixtures and from the scatter of the Siple results around a smooth line. The accuracy of the calibration, as determined from analysing international isotope standards, is better than $\pm 0.10\%$.

The $\delta^{13}\text{C}$ value and the concentration of CO_2 in the Siple ice samples are presented in Table 1 and Fig. 1. For comparison, the average values for air extracted from an interval of a South Pole ice core which is between 430 and 770 years old⁶, and results from direct atmospheric measurements at Mauna Loa (refs 8–10 and C. D. Keeling, personal communication) are also

Table 1 Analytical results of the Siple ice core

Average depth (m)	Gas age (yr AD)	Concentration (p.p.m.)	$\delta^{13}\text{C}$ (‰, PDB)
187.70	1744	276.8	-6.48
177.50	1764	276.7	-6.31
168.30	1791	279.7	-6.43
154.89	1816	283.8	-6.39
142.75	1839	283.1	-6.62
140.75	1843	287.4	-6.54
138.20	1847	286.8	-6.51
134.47	1854	288.2	-6.48
126.80	1869	289.3	-6.55
123.80	1874	289.5	-6.43
121.80	1878	290.3	-6.51
116.82	1887	292.3	-6.42
110.20	1899	295.8	-6.77
108.80	1903	294.8	-6.80
107.20	1905	296.9	-6.71
105.25	1909	299.2	-6.54
101.80	1915	300.5	-6.86
98.80	1921	301.6	-6.74
95.17	1927	305.5	-6.78
90.77	1935	306.6	-6.91
86.80	1943	307.9	-6.82
81.22	1953	312.7	-6.85

shown in Fig. 1. The average $\delta^{13}\text{C}$ value of the three samples from before 1800 (-6.41%) is in excellent agreement with the mean value for the South Pole samples ($-6.44 \pm 0.03\%$). This strengthens our confidence in the reliability of the experimental method. We cannot exclude the possibility that the isotopic composition of CO_2 trapped in ice is shifted compared to the atmospheric value. However, the agreement indicates that this is not the case, because isotopic fractionation processes active during trapping or storage of the air bubbles in ice could be expected to have different effects on the two cores which sample ice accumulated at different temperatures (the mean annual temperature is -24°C at Siple Station, and -51°C at the South Pole).

During the nineteenth and twentieth centuries, $\delta^{13}\text{C}$ slowly decreased, approximately antiparallel to the CO_2 concentration trend. The youngest ice core data fit well the direct atmospheric observation in 1956 of -6.79% (results given in ref. 8 and N_2O -corrected according to Mook and van der Hoek⁷). The overall decrease from the end of the eighteenth century to 1980 (-7.55% at Mauna Loa⁹) is -1.14% . The estimated overall uncertainty of this figure is $\pm 0.15\%$, including a possible difference in calibration between the two data sets.

The $\delta^{13}\text{C}$ record from the Siple ice core exhibits a step-like shift before 1900 and a plateau between ~ 1920 and 1950. There is also a less well pronounced plateau in the CO_2 results at the same depth. It is premature at this stage to decide whether this actually reflects an accelerated atmospheric CO_2 and $\delta^{13}\text{C}$ change around the turn of the twentieth century, or whether it represents an artefact.

The ice-core results can be compared with tree-ring $\delta^{13}\text{C}$ records, measured by a number of authors^{11,12}. Tree-ring $\delta^{13}\text{C}$ values generally exhibit considerable scatter, due to the fact that the $^{13}\text{C}/^{12}\text{C}$ fractionation varies from plant to plant as well as from year-ring to year-ring within a plant. Different records, each an average over a number of trees, have yielded changes from 1800 to 1980 ranging from 0% (ref. 11) to -2% (ref. 12). Thus, it is difficult to compare those results with ours. Stuiver *et al.*¹³ attempted to correct for variable physiological isotope fractionation by making use of a correlation between $\delta^{13}\text{C}$ and tree-ring area, and obtained a change of -0.8% from 1800 to 1980. Interestingly, many trees show constant (or even increasing) values in the period 1920 to 1950^{12,14}, which is also indicated by our results.

The CO_2 results (Fig. 1b) include, besides the IR laser absorption data of Neftel *et al.*¹, additional results obtained by gas

chromatography. They indicate a pre-industrial (pre-1800) concentration of ~ 280 p.p.m., which is nearly equal, within a few p.p.m., to the average for the South Pole samples (283 ± 5 p.p.m., age interval 430–770 years). We therefore have no reason to assume any secular fluctuation exceeding 10 p.p.m. between the years of 1200–1800.

In order to compare the $\delta^{13}\text{C}$ and the CO_2 concentration results and to interpret these results in terms of CO_2 release rates into the atmosphere, it was necessary to use a model of the global carbon cycle^{15,16}. Within the framework of the carbon cycle model, the CO_2 and the $\delta^{13}\text{C}$ records from the Siple ice core are consistent; the calculated $\delta^{13}\text{C}$ trend (using the CO_2 concentration history of Fig. 1b as model input) agrees well with the observations shown in Fig. 1a. Taking the Siple CO_2 and $\delta^{13}\text{C}$ records as model inputs, we calculated the release rates of CO_2 into the atmosphere. The results depend on the structure of the oceanic and biospheric sub-models, but indicate that the cumulative non-fossil CO_2 production until 1980 was about half of the cumulative fossil production, and 50% or more of this non-fossil CO_2 was released prior to 1900.

What additional information can be obtained from isotopic data that is not already contained in the CO_2 concentration record? Fossil and biospheric carbon both have very similar $\delta^{13}\text{C}$ values, but oceanic carbon is considerably richer in ^{13}C . Thus, $\delta^{13}\text{C}$ can, in principle, be used to decide whether the non-fossil CO_2 input, as reconstructed from the concentration record, was of biospheric or of oceanic origin. CO_2 escaping from the ocean is isotopically very similar to atmospheric carbon dioxide, so that input from the ocean would not have a significant effect on $\delta^{13}\text{C}$ values in contrast to input from the biosphere. The difference in atmospheric $\delta^{13}\text{C}$ values between a biospheric and an oceanic CO_2 input depends on the input history. As will be discussed in detail elsewhere¹⁷, model calculations with the non-fossil input history back-calculated from the CO_2 record of Fig. 1b yield a relatively small difference between biospheric and oceanic sources in the overall isotopic decrease. This result can be understood by considering the different behaviour of CO_2 concentration and isotopic composition with respect to atmosphere-ocean interaction. The oceanic uptake capacity for CO_2 is limited compared to that for isotopic perturbations by a chemical buffering mechanism, and an isotopic perturbation is additionally attenuated by exchange with the biosphere. Thus, a large CO_2 release during the 19th century still significantly affects the present-day atmospheric CO_2 concentration, but not the $\delta^{13}\text{C}$ value. In other words, $\delta^{13}\text{C}$ values integrate over a relatively short time and therefore, a $\delta^{13}\text{C}$ record can reveal the time variations of the CO_2 release rates in greater detail than a record of CO_2 concentrations. For this reason, further ice-core measurements are required to establish the past atmospheric history of CO_2 concentration and of $\delta^{13}\text{C}$ values.

We thank K. Hänni for help with mass-spectrometric measurements and CO_2 extraction. Supported by the Swiss NSF.

Received 23 June; accepted 13 August 1986.

- Neftel, A., Moor, E., Oeschger, H. & Stauffer, B. *Nature* **315**, 45–47 (1985).
- Stauffer, B., Fischer, G., Neftel, A. & Oeschger, H. *Science* **229**, 1386–1388 (1985).
- Schwander, J. & Stauffer, B. *Nature* **311**, 45–47 (1984).
- Schwander, J. thesis, Univ. Bern (1984).
- Moor, E. & Stauffer, B. *J. Glaciol.* **30**, 358 (1984).
- Friedli, H., Moor, E., Oeschger, H., Siegenthaler, U. & Stauffer, B. *Geophys. Res. Lett.* **11**, 1145–1148 (1984).
- Mook, W. G. & van der Hoek, S. *Isotopic Geosci.* **1**, 237–242 (1983).
- Keeling, C. D., Mook, W. G. & Tans, P. P. *Nature* **277**, 121–123 (1979).
- Mook, W. G., Koopmans, M., Carter, A. F. & Keeling, C. D. *J. geophys. Res.* **88**, 10915–10933 (1983).
- Heimann, M., Keeling, C. D. & Mook, W. G. *Tellus* (in the press).
- Francey, R. J. *Nature* **290**, 232–235 (1981).
- Freyer, H. D. & Belacy, N. *J. geophys. Res.* **88**, 6844–6852 (1983).
- Stuiver, M., Burk, R. L. & Quay, P. D. *J. geophys. Res.* **88**, 11731–11748 (1984).
- Tans, P. P. & Mook, W. G. *Tellus* **32**, 268–283 (1980).
- Siegenthaler, U., Heimann, M. & Oeschger, H. in *Carbon Dioxide, Climate and Society* (ed. Williams, J.) 79–87 (Pergamon, Oxford, 1981).
- Peng, T. H., Broecker, W. S., Freyer, H. D. & Trumbore, S. *J. geophys. Res.* **88**, 3609–3620 (1983).
- Siegenthaler, U. & Oeschger, H. *Tellus* **39B** (in the press).
- Craig, H. *Geochim. cosmochim. Acta* **3**, 53–92 (1953).

Seasonal variation of methylarsenic compounds in airborne particulate matter

Hitoshi Mukai*, Yoshinari Ambe*, Tatsunori Muku†, Kazuo Takeshita† & Tsuneo Fukuma†

* Chemistry and Physics Division, National Institute for Environmental Studies, Yatabe Tsukuba, Ibaraki, 305 Japan

† Saigo Health Center, Minato Saigo, Oki, Simane 685 Japan

Inorganic and organic arsenic compounds in nature are methylated by biological action and released into the air, forming trimethylarsine and dimethylarsine¹⁻⁵. These compounds are oxidized in the air to dimethylarsinic acid and trimethylarsineoxide⁶ and exist as particles^{7,8}. In this study, we examine seasonal variations of methylarsenic compounds in airborne particulate matter to try to understand the characteristics of the methylation of arsenic in nature. In summer a high concentration of dimethyl and trimethyl forms of arsenic was observed, while in winter the levels were very low. This seasonal variation can be attributed to the different temperatures of the environment during these seasons. The activation energy of methylation of arsenic was estimated to be $\sim 12 \text{ kcal mol}^{-1}$ from Arrhenius plots and can be used to predict the behaviour of these methylarsenic compounds, which are considered to form an important portion of the arsenic present in airborne particulate matter.

The natural methylation of arsenic has been reported in lakes, seas and soils⁹⁻¹¹. Challenger and Higginbottom² reported that bread mould produced trimethylarsine from inorganic arsenic. Cox and Alexander⁴ also showed that fungi in sewage could produce trimethylarsine from inorganic and methylarsenic compounds. Dimethylarsine is synthesized by methanobacterium³. These arsines are released from various environmental media and are present in air. Johnson and Braman⁷ detected both gaseous and particulate trimethyl and dimethyl forms of arsenic in the atmosphere. The methylation of arsenic is thus important for determining the mobility of this element in nature. However, biological methylation in nature is not well understood because the concentrations of these methylated compounds in the air

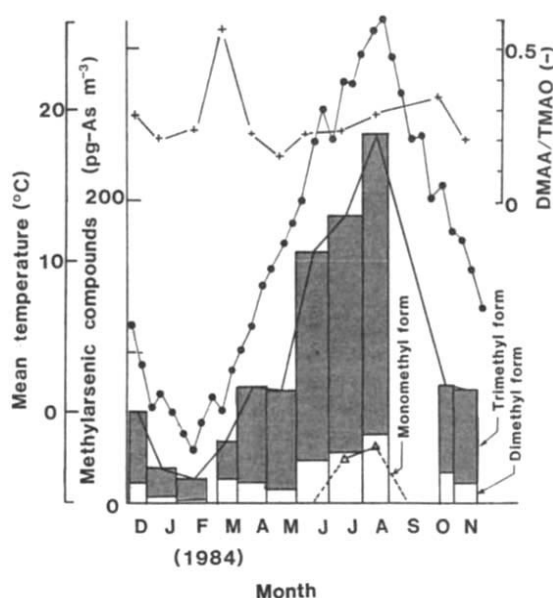


Fig. 1 Seasonal variations of methylarsenic compounds at Oki Islands, Japan. Δ , Monomethyl forms of arsenic. +, The ratio of dimethyl form (pg-As m^{-3}) to trimethyl form (pg-As m^{-3}). $\bullet$, Mean temperature of every 10 days.

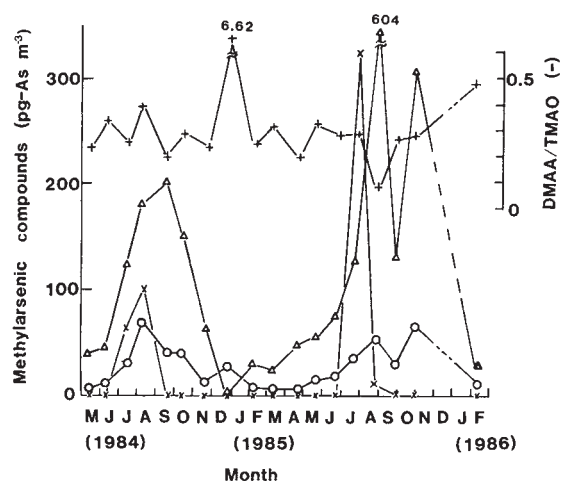


Fig. 2 Seasonal variations of methylarsenic compounds at Tsukuba, Japan. Δ , Concentration trimethyl; $\circ$, dimethyl; $\times$, monomethyl forms of arsenic. +, The ratio of dimethyl form to trimethyl form.

are very low. Methylarsenic compounds, for example, are oxidized in the air and converted to particulate form (trimethylarsineoxide and dimethylarsinic acid)⁶. Therefore, concentrations of these compounds in airborne particulate matter reflect the activity of biomethylation in nature.

Sampling of airborne particulate matter was carried out over 1 year at two sites in Japan, Oki Islands (an unpolluted area) in the Japan Sea and Tsukuba (a rural inland area). These two sampling sites are $\sim 1,000 \text{ km}$ apart. The sampling at Oki Islands was carried out over 1 month with a low volume air sampler at the top of a mountain (200 m) (December 1983–November 1984). The samples from Tsukuba were collected with a high-volume air sampler for 3 days each month (May 1984–February 1986). Quartz fibre filters (PALLFLEX 2500QAST) were used; loaded filters were preserved in a store room at -20°C .

The hydride generation method proposed by Braman *et al.*^{7,12} was adapted to analyse the methylarsenic compounds in airborne particulate matter. Dimethyl and trimethyl forms of arsenic (possibly dimethylarsinic acid and trimethylarsineoxide) were extracted from the filter with 10 cm^3 of $1 \text{ M HCl} + 0.02 \text{ M EDTA}$ (9:1) and reduced with 6 cm^3 of $10\% \text{ NaBH}_4$ in 1 M HCl containing $20 \mu\text{mol EDTA}$. The arsines produced were trapped in a half-packed U-tube dipped in liquid nitrogen. Following gas chromatography, each arsine separated was detected as arsenic by atomic absorption spectroscopy. The details of this method are described elsewhere¹³. The detection limits of monomethyl, dimethyl and trimethyl forms of arsenic by this method were $\sim 2 \text{ pg-As m}^{-3}$.

Figure 1 shows the variations of methylarsenic compounds in airborne particulate matter sampled at Oki Islands. High concentrations of dimethyl and trimethyl forms were observed in summer. The pattern of the seasonal changes of the concentrations of dimethyl and trimethyl forms was found to correlate with the mean temperature. The order of these concentrations was similar to that reported by Johnson and Braman⁷. The ratio of the dimethyl form to trimethyl form was 0.15–0.34 in most cases, indicating that trimethylarsine was the dominant form of methylated arsenic produced by biomethylation. Traces of the monomethyl form were detected in the July and August samples.

Although the ratio of these methylarsenic compounds to inorganic arsenic in airborne particulate matter was $<10\%$, the contribution of methylated arsenic cannot be ignored, because the concentration of methylarsenic compounds in airborne particulate matter often amounted to 18 p.p.m. whereas the concentration of arsenic in the Earth's crust is 2 p.p.m..

Figure 2 illustrates the seasonal variations of methylarsenic compounds at Tsukuba. Although the two sampling sites are

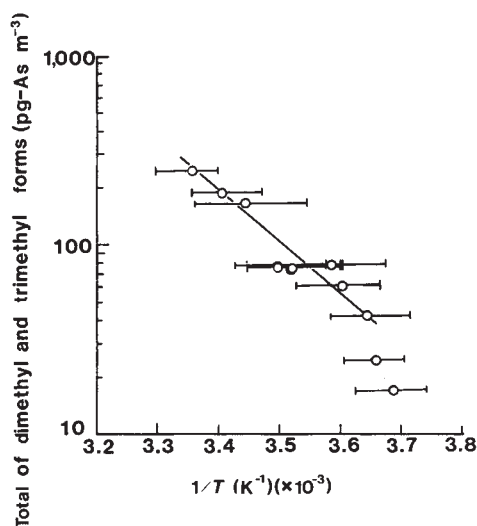


Fig. 3 Arrhenius plot using the total concentration of the dimethyl and trimethyl arsenic compounds in the air (particulate form) at Oki Islands. The line was drawn using nine points, and disregards two winter points, when snow covered the land surface. The horizontal line of each point indicates the range between the mean value of maximum and minimum temperature.

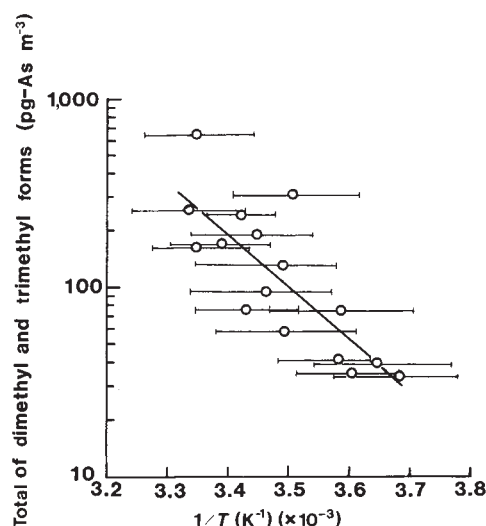


Fig. 4 Relation between the total of the concentrations of the dimethyl and trimethyl arsenic compounds (particulate form) at Tsukuba. The line obtained from Fig. 3 is shown. The horizontal line at each point indicates the range between maximum and minimum temperature during the sampling period.

far apart from each other, the patterns of the variation and the concentration of the methylarsenic compounds were very similar in Oki Islands and Tsukuba. At both locations the concentrations of dimethyl and trimethyl forms were high in summer, but low in winter. The monomethyl form, which was detected in the summer samples at Oki Islands, was also detected only in the July and August samples, and was considered to be caused by contamination with a pesticide¹⁴ (iron methane arsonate) spread over rice fields 20–30 km away from the sampling site, not biomethylation. Therefore, we shall not consider this form in the following discussion.

The seasonal variations of the dimethyl and trimethyl forms of arsenic can be attributed to the change of temperature of the environment. Tanaka *et al.*⁸ reported that a high concentration of dimethylarsenic compound in airborne particulate matter was observed during the summer. In the case of methylation of selenium, Chau *et al.*¹⁵ pointed that the production of volatile selenium was also temperature dependent. Figure 3 shows the relation between the total dimethylarsenic and trimethylarsenic concentrations (mean concentration for ~1 month) at Oki Islands and the reciprocal of the mean temperature of the sampling period. If we assume that arsines diffuse into the atmosphere faster than their rate of production, and are oxidized constantly, the concentration of these compounds is considered to reflect the rate of production of methylarsenic compounds, and the activation energy of biomethylation of arsenic in nature is estimated to be 12 kcal mol⁻¹ from the slope of the graph (avoiding the two winter samples; using all the points gives a value of ~16 kcal mol⁻¹). This value is consistent with those of general biological reactions (8–18 kcal mol⁻¹). But, as biomethylation is influenced not only by temperature but also by other factors such as humidity^{11,16} and the concentration of these compounds in air is affected by meteorological conditions (such as mixing depth, wind speed and weather), the activation energy obtained is considered to be the apparent value. It will reflect various environmental factors, even though temperature is a major factor. For example, if the rate that the arsine gas is oxidized and converted to the airborne particulate matter is faster in summer, the slope of the line of the Arrhenius plot will be overestimated compared with the real value, whereas the existence of a surface inversion layer in winter might make the concentration of airborne particulate matter relatively higher

and lead to an underestimation of the activation energy. The effect of an inversion layer may be minor, because the sampling site is located on top of the mountain (200 m). Further study will be necessary in order to understand the conversion rate, which includes the rate of heterogeneous reaction between arsine gas and airborne particulate matter as well as the rate of homogeneous oxidation reaction with oxygen⁶, peroxides and hydroxyl radicals.

Figure 4 shows the data from Tsukuba. Many points were distributed along the line drawn from the data in Fig. 3, indicating that a similar activation energy can be obtained even though the sampling sites are different. The deviation of two points from the line suggests that the data obtained during short-period sampling is strongly influenced by the meteorological conditions at the time and that the line does not describe the maximum concentration corresponding to each temperature.

Our data showed that the methylation of arsenic in nature is an important source of arsenic in the air and is influenced by temperature. The activation energy estimated above can be used to predict the behaviour of methylarsenic compounds in the air throughout the year, although it is not known to which methylation reaction this value corresponds. These findings do not give any new insight into biological methylation; however, they give important information concerning the contribution of biomethylation to the geological cycles of trace elements (As, Se and so on). The data shown here were only from the particulate form; the gaseous form of trimethylarsine needs further investigation. The measurements of the change of the rate of the production of methylarsenic compounds with the temperature of variables such as the land surface, which is one of the most probable sources of methylarsenic compounds, are also needed for comparison with the obtained activation energy.

We thank members of Saigo Health Center and NTT (Saigo) for the field assistance; Dr T. Uehiro, Dr M. Morita and Dr Y. Yokouchi for useful advice.

Received 9 June; accepted 3 September 1986.

1. Challenger, F., Higginbottom, C. & Ellis, L. *J. chem. Soc.* 95–101 (1933).
2. Challenger, F. & Higginbottom, C. *Biochem. J.* 29, 1757–1788 (1935).
3. McBride, B. C. & Wolfe, R. S. *Biochemistry* 10, 4312–4317 (1971).
4. Cox, D. P. & Alexander, M. *Bull. Envir. Contam. Toxicol.* 9, 84–88 (1973).
5. Wood, J. M. *Science* 183, 1049–1052 (1974).
6. Parris, G. E. & Brinckman, F. E. *Envir. Sci. Technol.* 10, 1128–1134 (1976).

7. Johnson, D. L. & Brame, R. S. *Chemosphere* **4**, 333-338 (1975).
8. Tanaka, S., Kaneko, M. & Hashimoto, Y. *J. chem. Soc. Jap.* **637-642** (1984).
9. Brame, R. S. & Foreback, C. C. *Science* **182**, 1247-1249 (1973).
10. Andreae, M. O. *Limnol. Oceanogr.* **24**, 440-452 (1979).
11. Akins, M. B. & Lewis, R. J. *Soil Sci. Soc. Am. J.* **40**, 655-658 (1976).
12. Brame, R. S., Johnson, D. L., Foreback, C. C., Ammons, J. M. & Bricker, J. L. *Analyt. Chem.* **49**, 621-625 (1977).
13. Mukai, H. & Ambe, Y. *Analyt. chim. Acta.* (submitted).
14. Mukai, H. & Ambe, Y. *Atmos. Envir.* (in the press).
15. Chau, Y. K., Wong, P. T. S., Silverberg, B. A., Luxon, P. L. & Bengert, G. A. *Science* **192**, 1130-1131 (1976).
16. Woolson, E. A. *Envir. Hlth Perspect.* **19**, 73-81 (1977).

Reversed magnetization in pyroclastics from the 1985 eruption of Nevado del Ruiz, Colombia

Friedrich Heller*, Juan Carlos Carracedo† & Vicente Soler†

* Institut für Geophysik, ETH-Hönggerberg, CH-8093 Zürich, Switzerland

† Estación Volcanológica de Canarias, Instituto de Recursos Naturales, (CSIC), La Laguna, Tenerife, Canary Islands, Spain

One of the fundamental principles of palaeomagnetism and rock magnetism states that the direction of natural remanent magnetization (NRM) is acquired parallel to the applied field. However, ferromagnetic mineral phases with complex microstructures may violate this rule^{1,2} and produce a magnetization vector which is antiparallel to the direction of the external field. Andesitic pumice, which was hurled several kilometres during the disastrous 1985 eruption of the Nevado del Ruiz volcano (Colombia), carries a stable but reversed NRM with southerly declination and negative inclination. Heating experiments show that this magnetization is due to a self-reversal mechanism which also induces a reversed thermoremanent magnetization (TRM) in the laboratory field. Thus we have used the Nevado del Ruiz 1985 pumice, which was magnetized in the present well-known geomagnetic field, to demonstrate for the first time that self reversal has actually controlled the process of NRM acquisition.

The Nevado del Ruiz, a stratovolcano of the Ruiz-Tolima complex in the northernmost part of the Andean Cordillera (lat. 4°53' N, long. 75°10' W, altitude 5,400 m, regional geomagnetic field inclination +31°), covered by a 17 km² ice cap, has had two major eruptions in 1595 and 1845, and its last explosive eruption was on 13 November 1985 (ref. 3). Lithic fragments, ash, and juvenile scoria and pumice fragments were emitted during the last eruption and an eruptive column up to ~5,600 m in height was formed. Most of the hot juvenile eruptive products fell on the ice cap. The catastrophic mudflow thus generated travelled down the radial river valleys causing the death of 25,000 people and damage assessed at US\$ 300 million.

Dacitic andesite pumice fragments were spread around the emission centre in the form of pyroclastic falls. Some of the fragments were embedded in the mud as they fell, or penetrated cushions of the *Plantago rigida* vegetation. Where they lodged, moderate thermal effects were observed in smashed plant fragments.

To investigate the temperature of deposition of this pumice, hand samples of encrusted fragments (5-10 cm in diameter), oriented with a magnetic compass, were collected within a radius of 5-6 km from the emission centre. The oriented samples were collected over a 3-week period after the 13 November eruption by one of us (J.C.C.), who formed part of the Scientific Commission sent by the Spanish Ministry of Foreign Affairs to collaborate in the study and surveillance of the Nevado del Ruiz volcano.

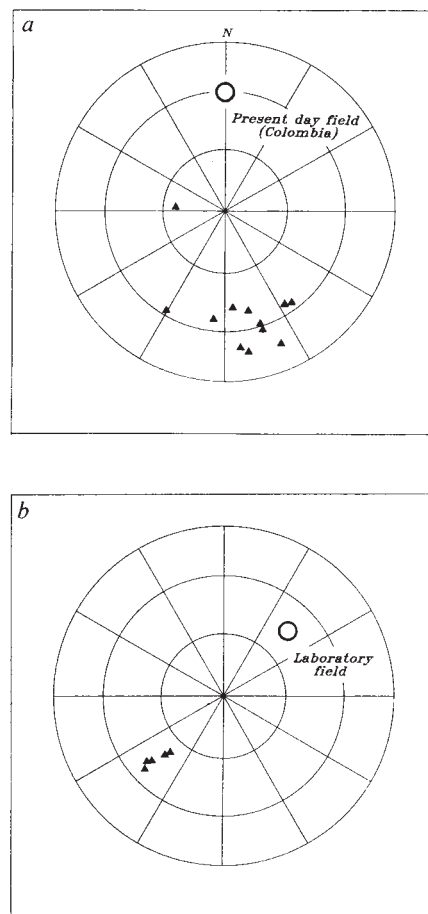


Fig. 1 Directions of: a, natural remanent magnetization (NRM); mean direction: declination = 170.7°; inclination = -35.9°; α_{95} = 14.2; 12 samples plotted. b, Laboratory-induced thermoremanent magnetization (TRM); mean direction: declination = 226.6°; inclination = -44.7°; α_{95} = 6.3; 5 samples plotted. Plots are stereographic projections together with applied field directions. Field and magnetization vectors are antiparallel. ○, Lower; ▲, upper hemisphere.

The result of the initial magnetic remanence measurements is shown in Fig. 1. All samples carry a reversed NRM which—within statistical error—is antiparallel to the present field of the region. Alternating field (AF) demagnetization reduces the directional scatter that results from small orientation inaccuracies and from additional magnetization components of normal polarity and of spurious direction. The main reversed NRM component is stable up to 80 mT and has median destructive fields between 10 mT and 30 mT.

When the NRM vector is measured during repeated heating and cooling cycles at variable temperatures in air and in zero field⁴, a polarity change is registered at temperatures ~120 °C (Fig. 2), the strong reversed room temperature component being completely removed at 150-160 °C. Above these temperatures the NRM polarity is always normal (parallel to the present field). When cooling back to room temperature after the first heating cycle, the reversed component develops again in several samples (Fig. 2a). Thus, the reversal mechanism is still active in these laboratory conditions. On heating to higher temperatures the normal NRM component disappears between 350 and 400 °C.

Two optically homogeneous magnetic mineral phases of different colour and reflectivity are observed using the ore microscope. A brownish, low reflectivity but strongly magnetic phase is attributed to titanomagnetite. A whitish-yellow, higher reflectivity phase is attributed to magnetite.

Geochemical correlation between southern African kimberlites and South Atlantic hotspots

Anton P. le Roex

Department of Geology, University of Cape Town, Rondebosch, 7700 South Africa

Recent isotopic and trace element data^{1,2} have led to the subdivision of southern African Cretaceous/Jurassic kimberlites into group I and group II varieties. The former have Sr, Nd and Pb isotopic signatures similar to those commonly associated with ocean island basalts and are thought to derive from within the asthenosphere¹. The high Sr and low Nd and ²⁰⁶Pb/²⁰⁴Pb isotopic ratios of group II kimberlites have been cited as evidence that these kimberlites are derived from a source within the sub-continental lithosphere^{1,3}. Basalts erupted at the surface expression of South Atlantic hotspots preserve isotopic and trace element ratios similar to those of group I and group II kimberlites. Hotspots falling within the Dupal anomaly⁴ have similar characteristics to group II kimberlites and those outside the Dupal anomaly to group I kimberlites. This correlation raises the possibility that the source region of southern African group II kimberlites is also located within the asthenosphere (within recycled lithospheric material) and not necessarily within the sub-continental lithosphere.

Southern African Cretaceous/Jurassic kimberlites comprise two dominant petrographic and geochemical varieties, namely so-called micaceous (group II) and basaltic or non-micaceous (group I) kimberlites^{1,5}. Both types contain diamonds, appear to be distributed throughout the Kaapvaal craton and show a broad age correlation with group I kimberlites (80–114 Myr), being generally younger than group II kimberlites (114–200 Myr)^{1,6}. Group I kimberlites have significantly lower ⁸⁷Sr/⁸⁶Sr and Rb/Sr ratios and higher ¹⁴³Nd/¹⁴⁴Nd, ²⁰⁶Pb/²⁰⁴Pb and U/Pb ratios than group II kimberlites^{1–3,5,7}.

Smith¹ has argued that these geochemical differences are not a result of alteration and/or crustal contamination but reflect substantial chemical differences in their mantle source regions. The slightly depleted source region signature of group I kimberlites, relative to bulk Earth, has been cited as evidence for an asthenospheric origin (their isotopic characteristics are similar to most ocean island basalts). In contrast, the enriched source region signature of group II kimberlites has been used to argue for an origin within the ancient enriched sub-continental lithosphere^{1–3}.

Hart⁴ has recently delineated a major region between southern Africa and South America that forms part of his proposed Dupal anomaly, in which the isotopic signature of hotspot lavas is distinct from those occurring to the north of ~20° S and to the south of ~43° S. More recent data (S. R. Hart, unpublished data) shows that in the South Atlantic this anomaly extends as far south as the recently documented Shona hotspot at ~54° S^{8,9}. Basalts that have erupted at hotspots within the Dupal region have Sr and Nd isotopic ratios enriched compared with the bulk Earth and have higher ²⁰⁷Pb/²⁰⁴Pb and ²⁰⁸Pb/²⁰⁴Pb ratios than most other oceanic volcanics. Here I compare the isotopic and trace element 'signatures' of South Atlantic hotspots and southern African kimberlites.

Basalts from the South Atlantic describe two distinct trends in terms of Zr/Nb ratio versus ⁸⁷Sr/⁸⁶Sr and ¹⁴³Nd/¹⁴⁴Nd ratios (Fig. 1). The more radiogenic Sr and less radiogenic Nd trends for a given Zr/Nb ratio correspond to basalts erupted within the Dupal anomaly and comprise the hotspots of Gough, Discovery and Tristan da Cunha and MORB in the vicinity of the Shona hotspot and, by implication, the as yet unsampled Shona hotspot⁸. The lower trend in each diagram includes the hotspots of Bouvet, Marion and Ascension and MORB from

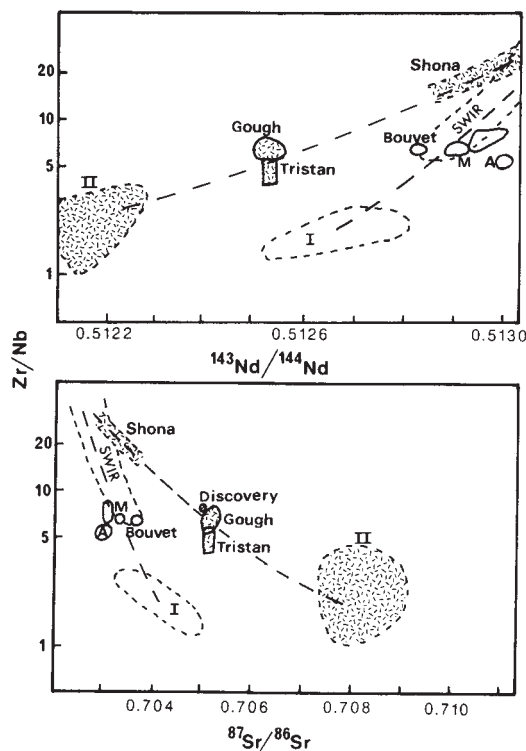


Fig. 1 Correlation between Zr/Nb ratio and ⁸⁷Sr/⁸⁶Sr and ¹⁴³Nd/¹⁴⁴Nd ratios in South Atlantic basalts and group I and group II kimberlites. Initial isotope ratios of the kimberlites have been used for plotting. SWIR, Southwest Indian Ridge; II, group II kimberlites; I, group I kimberlites; A, Ascension; M, Marion (two fields). Sources of data: refs 1, 3, 8, 19, 25–28, 31, 34, S. R. Hart (unpublished data) and A.P. le R. (unpublished data).

the vicinity of the Bouvet hotspot. It is evident from Fig. 1 that group I kimberlites plot on the extension of the 'Bouvet' trend whereas group II kimberlites plot on the extension of the Dupal trend.

A similar correlation exists for incompatible trace element variations. Figure 2 shows the variation of Ba/Nb, Zr/Nb and La/Nb ratios in MORB, South Atlantic hotspot basalts and group I and group II kimberlites. Hotspots within the Dupal anomaly (that is, Gough, Discovery, Tristan da Cunha and the associated Walvis Ridge and Rio Grande Rise) are characterized by high Ba/Nb (8–22), La/Nb (0.75–1.25) and low Zr/Nb (<10) ratios and define a trend extending away from the field of Bouvet, Marion and other non-Dupal-area hotspots towards group II kimberlites (Ba/Nb > 15, La/Nb > 1, Zr/Nb < 3). Although data are not yet available from the Shona hotspot, the high Ba/Nb ratios of associated southern Mid-Atlantic Ridge enriched MORB (Fig. 2) indicates that this hotspot is also likely to be characterized by high Ba/Nb ratios. Group I kimberlites, on the other hand, have low Ba/Nb (<8), La/Nb (<0.8) and Zr/Nb (<4) ratios, similar to those of the non-Dupal hotspots Bouvet, Marion and Ascension (Ba/Nb < 9, La/Nb < 0.6 and Zr/Nb < 10). Depleted MORB in these diagrams defines a very distinct trend extending towards high Zr/Nb (>20) and La/Nb (>1.0) ratios at low (<7) Ba/Nb ratios.

Figure 3 extends this correlation between group II kimberlites and South Atlantic Dupal hotspot volcanics to include Pb isotopes. It is evident that these lavas all plot on a sub-parallel Pb isotope array characterized by low ²⁰⁶Pb/²⁰⁴Pb and high ²⁰⁸Pb/²⁰⁴Pb and (excluding Tristan da Cunha) ²⁰⁷Pb/²⁰⁴Pb ratios relative to MORB, non-Dupal-area hotspot volcanics and group I kimberlites. The high ²⁰⁸Pb and ²⁰⁷Pb of enriched southern Mid-Atlantic Ridge basalts (Fig. 3; S. R. Hart, unpublished data) associated with the Shona hotspot indicate that the latter is likely to have Pb isotopic signature similar to that of Gough.

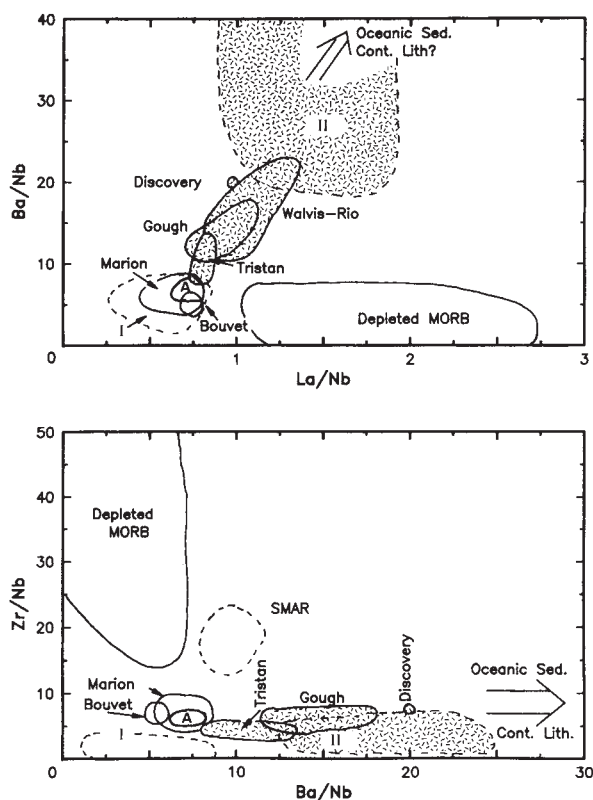


Fig. 2 Covariation of Ba/Nb, Zr/Nb and La/Nb ratios in South Atlantic basalts and group I and group II kimberlites. Compositions of oceanic sediment¹⁹ and sub-continental lithosphere²⁶ are shown for comparison. Walvis-Rio, Walvis Ridge and Rio Grande Rise; SMAR, southern Mid-Atlantic Ridge; other abbreviations as in Fig. 1. Data from refs 3, 10, 19, 25, 26, 31-34 and A.P. le R. (unpublished data).

Zr/Nb, Ba/Nb and La/Nb are highly incompatible trace element ratios that are not normally fractionated during partial melting processes. However, the very low degrees of partial melting possibly associated with alkaline, and particularly kimberlite, magma generation require that some caution be exercised when interpreting variations in such inter-element ratios. Nevertheless, the good correlations exhibited between these ratios and Sr, Nd and Pb isotopic ratios suggest that the variations shown in Figs 1 and 2 reflect source region composition and are not primarily due to variations in degree of partial melting or residual source mineralogy.

Although the trends depicted in Fig. 1 correspond broadly to mixing between a depleted MORB-type reservoir and highly enriched reservoirs (reflected by group I and group II kimberlites as possible end-members), Figs 2 and 3 negate any simple mixing relation between the Dupal hotspot-group II kimberlite trend and a depleted MORB-type source. Even within the Dupal hotspot-group II kimberlite trend, relations are complex and there is, for example, no simple correlation between $^{87}\text{Sr}/^{86}\text{Sr}$ and $^{206}\text{Pb}/^{204}\text{Pb}$ along the Pb-Pb arrays⁴.

Before evaluating possible implications of the geochemical correlations shown in Figs 1-3, it is necessary to consider current models for the origin of hotspot volcanism. Allègre and Turcotte¹¹ recently summarized current ideas on hotspot source regions and have proposed an integrated model which recognizes two 'end-member' hotspot types. The first derives from upwelling of primordial or near-primordial mantle across the asthenosphere-mesosphere boundary, the second from upwelling from a 'mesosphere boundary layer' of recycled (subducted) oceanic lithosphere including pelagic sediments^{12,13} and/or delaminated sub-continental lithosphere¹⁴. Delaminated litho-

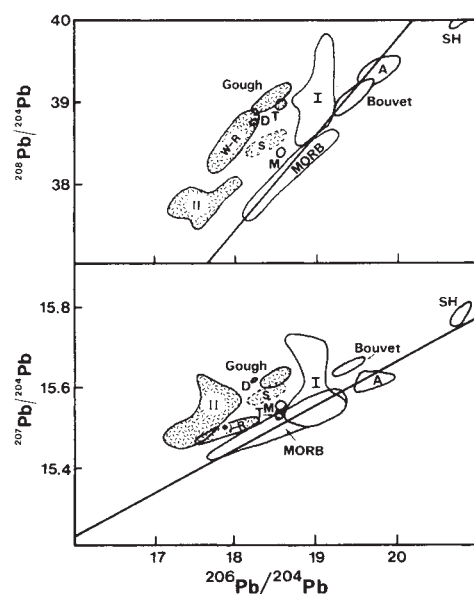


Fig. 3 Pb isotope variation in South Atlantic basalts, group I and group II kimberlites and MORB. Initial isotopic ratios of the kimberlites have been used for plotting. NHRL, Northern Hemisphere reference line³; D, Discovery; T, Tristan; SH, St Helena; other abbreviations as in Figs 1 and 2. Data from refs 1, 2, 19, 21 and S. R. Hart (unpublished).

sphere may alternatively occur as 'rafts' within the asthenosphere rather than as a boundary layer¹⁵. Evidence for a primordial source component derives mainly from the high $^3\text{He}/^4\text{He}$ isotopic ratios associated with hotspots such as Hawaii and Iceland¹⁶, and evidence for a recycled component derives from the higher $^{208}\text{Pb}/^{204}\text{Pb}$ and $^{207}\text{Pb}/^{204}\text{Pb}$ ratios and low $^3\text{He}/^4\text{He}$ ratios of hotspots such as Gough and Tristan da Cunha^{11,16}. The sub-parallel Pb isotope array of Dupal hotspots also requires that the recycled material composing this mesospheric boundary layer is ancient (>1 Gyr)^{4,11}.

In the South Atlantic, Bouvet (with a high $^3\text{He}/^4\text{He}$ ratio¹⁷) can be regarded as an example of a hotspot with a significant primordial component in its source region¹⁸, whereas Gough, Tristan da Cunha, Discovery and Shona appear to be hotspots with a major recycled component in their source regions^{11,16,18,19}. This interpretation is also consistent with the trace element variations shown in Fig. 2, where Gough, Tristan and Discovery define a trend extending away from compositions typical for normal ocean island basalt (containing a major primordial component), such as Bouvet, Marion and Ascension, towards high Ba/Nb and La/Nb ratios characteristic of recycled oceanic lithosphere plus sediment or of sub-continental lithosphere^{20,21}.

Trace element and isotopic variations in South Atlantic hotspots therefore suggest that the Dupal anomaly reflects an area where ancient recycled material exists within the asthenosphere. Upwelling from this material has resulted in hotspot volcanism with an apparent lithospheric signature, whereas upwelling outside this region has led to hotspot volcanism with a more normal ocean island basalt signature.

It is remarkable that in all three figures (Figs 1-3), group I and group II kimberlites plot either at the end of, or are superimposed on, the non-Dupal hotspot and Dupal hotspot trends respectively. Group I and group II kimberlites therefore appear to be derived from source regions similar to the enriched components invoked respectively for non-Dupal and Dupal hotspots. In group II kimberlites and Dupal hotspots this component has distinctively low $^{206}\text{Pb}/^{204}\text{Pb}$ and high $^{208}\text{Pb}/^{204}\text{Pb}$, Ba/Nb and La/Nb ratios. In group I kimberlites and non-Dupal hotspots this component has high $^{206}\text{Pb}/^{204}\text{Pb}$ and low Ba/Nb and La/Nb ratios.

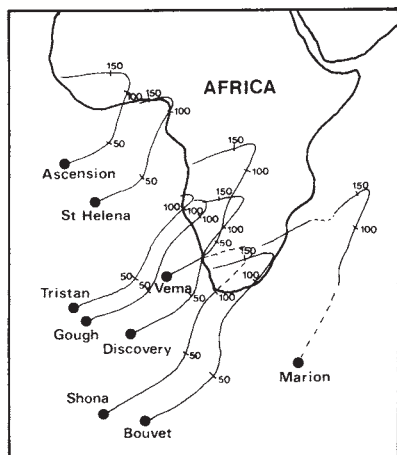


Fig. 4 Diagram modified from ref. 30 showing South Atlantic hotspots and their predicted palaeo-tracks. The Shona hotspot and associated track is from ref. 9. Numbers on tracks show ages in Myr.

Two separate source regions therefore occur within the asthenosphere beneath the South Atlantic capable of generating alkaline volcanism with essentially the same distinct geochemical signatures as shown by group I and group II kimberlites. We may extend this observation to recent plate-tectonic models advocating a correlation between the palaeo-position of hotspots and kimberlite magmatism²²⁻²⁴.

The African plate contains many hotspots, several of which appear to record the entire history of the opening of the South Atlantic (Fig. 4). On the basis of plate reconstructions, Cretaceous and Jurassic kimberlites in southern Africa have been explicitly related to the Bouvet and 'Meteor' (recently relocated and renamed Shona⁹) hotspots^{23,24}. Geochemical evidence presented here is consistent with such a model in that upwelling from these (or related) hotspots, when located beneath southern Africa (Fig. 4), could lead to the deep-seated melting and subsequent eruption of kimberlite magmas with inherently different geochemical signatures, reflecting those of their respective source regions. Kimberlites derived from upwelling associated with the Bouvet (or similar) hotspot would have group I geochemical characteristics, whereas those derived from upwelling associated with the Shona (or similar) hotspot would have group II characteristics. In this regard, note that kimberlites in Ghana, Guinea and Sierra Leone have group I characteristics, which is consistent with their location on the palaeo-track of the Ascension hotspot (J. J. Gurney, personal communication; Fig. 4), which has essentially the same trace element and isotopic signature as Bouvet (Figs 1-3).

Recent plate reconstructions^{9,30} are consistent with a relationship between the Shona hotspot and group II (114-150 Myr) kimberlites but are not readily compatible with a simple relationship between Cretaceous group I kimberlites and the Bouvet hotspot (Fig. 4). Group I kimberlites are too young to be directly associated with Morgan's³⁰ Bouvet palaeo-track and it is possible that in a more complex model, refinements to the plate-tectonic reconstructions, or the presence of an unrecognized non-Dupal hotspot located to the north-east of Bouvet will be required to resolve this problem.

The model proposed here places the source of both group I and group II kimberlites within the asthenosphere, with group II kimberlites deriving from recycled lithospheric material. This avoids having to derive spatially juxtaposed kimberlite magmas of similar age from two tectonically different source regions. Details relating specific hotspot tracks to specific kimberlite pipes still need to be resolved. However, the results presented here document the presence of a source region, located within the oceanic asthenosphere, capable of generating alkaline volcanism with essentially the same lithospheric signature as shown

by group II kimberlites, so lending credence to models whereby kimberlite magmatism in general is related to the palaeo-positions of postulated mantle hotspots.

I thank the University of Cape Town and the Foundation for Research Development for financial support; C. B. Smith for discussions; J. J. Gurney, A. M. Reid and S. R. Hart for critical comments; S. R. Hart for allowing me to use his unpublished data; and the British Museum for a sample of Discovery seamount.

Received 3 June; accepted 29 July 1986.

- Smith, C. B. *Nature* **304**, 51-54 (1983).
- Fraser, K. J., Hawkesworth, C. J., Erlank, A. J., Mitchell, R. H. & Scott-Smith, B. H. *Earth planet. Sci. Lett.* **76**, 57-70 (1985).
- Smith, C. B., Gurney, J. J., Skinner, E. M. W., Clement, C. R. & Ebrahim, N. *Trans. geol. Soc. S. Afr.* **88** (in the press).
- Hart, S. R. *Nature* **309**, 753-757 (1984).
- Dawson, J. B. in *Ultramafic and Related Rocks* (ed. Wyllie, P. J.), 241-251 (Wiley, New York, 1967).
- Smith, C. B., Allsopp, H. L., Kramers, J. D., Hutchinson, G. & Rodic, J. D. *Trans. geol. Soc. S. Afr.* **88** (in the press).
- Kramers, J. D. *Earth planet. Sci. Lett.* **34**, 419-431 (1977).
- le Roex, A. P., Dick, H. J. B., Reid, A. M., Gulen, L. & Erlank, A. J. *Geochim. cosmochim. Acta* (in the press).
- Hartnady, C. J. H. & le Roex, A. P. *Earth planet. Sci. Lett.* **75**, 245-257 (1985).
- Kempe, D. R. C. & Schilling, J.-G. *Contr. Miner. Petrol.* **44**, 101-115 (1974).
- Allègre, C. J. & Turcotte, D. L. *Geophys. Res. Lett.* **12**, 207-210 (1985).
- Hofmann, A. W. & White, W. M. *Earth planet. Sci. Lett.* **57**, 421-436 (1982).
- Ringwood, A. E. *J. Geol.* **90**, 611-643 (1982).
- McKenzie, D. & O'Nions, R. K. *Nature* **301**, 229-231 (1983).
- Hawkesworth, C. J., Mantovani, M., Taylor, P. N. & Palacz, Z. *Nature* **322**, 356-359 (1986).
- Kurz, M., Jenkins, W. J. & Hart, S. R. *Nature* **297**, 43-47 (1982).
- Kurz, M. thesis, Woods Hole Oceanogr. Inst. (1982).
- le Roex, A. P. in *Mantle Metasomatism* (eds Menzies, M. A. & Hawkesworth, C. J.) (Academic, London, in the press).
- le Roex, A. P. *J. Petrol.* **26**, 149-186 (1985).
- Hole, M. J., Saunders, A. D., Marriner, G. F. & Tarney, J. J. *geol. Soc. Lond.* **141**, 453-472 (1984).
- Hawkesworth, C. J., van Calsteren, P., Rogers, N. W. & Menzies, M. A. in *Mantle Metasomatism* (eds Menzies, M. A. & Hawkesworth, C. J.) (Academic, London, in the press).
- Duncan, R. A., Hargraves, R. B. & Brey, G. P. *Geol. Mag.* **115**, 317-327 (1978).
- Crough, S. T., Morgan, J. W. & Hargraves, R. B. *Earth planet. Sci. Lett.* **50**, 260-274 (1980).
- Duncan, R. A. *Tectonophysics* **74**, 29-42 (1981).
- le Roex, A. P. *et al. J. Petrol.* **24**, 267-318 (1983).
- le Roex, A. P. & Erlank, A. J. *J. Volc. geotherm. Res.* **13**, 309-338 (1982).
- O'Nions, R. K., Hamilton, P. J. & Evensen, N. M. *Earth planet. Sci. Lett.* **34**, 13-22 (1977).
- Sun, S.-S. *Phil. Trans. R. Soc. A297*, 409-445 (1980).
- Erlank, A. J. *et al. in Mantle Metasomatism* (eds Menzies, M. A. & Hawkesworth, C. J.) (Academic, London, in the press).
- Morgan, J. W. *Tectonophysics* **94**, 123-139 (1983).
- Kable, E. J. D. thesis, Univ. Cape Town (1972).
- Thompson, G., Humphris, S. & Schilling, J.-G. *Init. Rep. DSDP* **72**, 457-466 (1983).
- Humphris, S. & Thompson, G. *Earth planet. Sci. Lett.* **66**, 223-242 (1983).
- Harris, C., Bell, J. D. & Atkins, F. B. *Earth planet. Sci. Lett.* **60**, 79-85 (1982).

Archaean flow-top alteration zones formed initially in a low-temperature sulphate-rich environment

Donald R. Lowe & Gary R. Byerly

Department of Geology, Louisiana State University, Baton Rouge, Louisiana 70803, USA

The 3,500-3,400 Myr Onverwacht Group in the southern part of the Barberton Greenstone Belt, South Africa, includes up to 12 km of komatiitic and tholeiitic lavas containing minor sedimentary and felsic volcanic units¹. Komatiitic lavas showing spinifex textures, chrome spinels and high Ni and Cr contents are common throughout the sequence². The rocks have been subject to polyphase Archaean deformation and low-grade metasomatism²⁻⁴. At several positions within the Onverwacht Group (Fig. 1), the tops of komatiitic lavas underlying chert beds are marked regionally by 1-50-m thick zones of silicification, carbonatization and K₂O enrichment (Fig. 2). Because these zones have been variously interpreted to represent the felsic tops of mafic-to-felsic volcanic cycles, weathering horizons and shear zones, their origin is germane to any interpretation of the stratigraphy, structure and evolution of the Barberton belt. Here we present evidence that they formed in

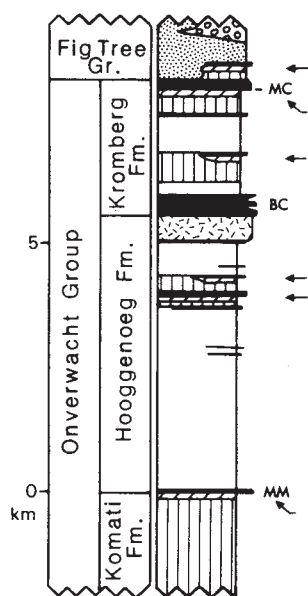


Fig. 1 Generalized stratigraphic column of the Onverwacht and lower Fig Tree Groups in the southern part of the Barberton Greenstone Belt showing the distribution of komatiitic flow top alteration zones (arrows). Symbols: vertical rule, komatiitic volcanic units; diagonal rule, flow-top alteration zones; unpatterned, tholeiitic units; random line segments, felsic units; black, cherts; stipple and circles, other sedimentary units; MM, Middle Marker; BC, Buck Ridge Chert and MC, Msauli Chert.

large part by low-temperature, near-surface alteration of komatiitic lavas during intervals of volcanic quiescence. They do not mark fault zones and cannot be used to subdivide the greenstone belt sequence into a series of fault-bounded structural slices.

The altered rock (Fig. 3a) consists of green, silicified komatiite cut by several generations of veins but, most distinctively, by anastomosing, crudely stratiform veins of silica and carbonate². Komatiitic remnants have been altered to microcrystalline aggregates that can include chlorite, sericite, chrome-rich sericite, talc, carbonate and chert, but many show well-preserved silica pseudomorphs of igneous spinifex and cumulate minerals². Of primary igneous minerals, only chrome-rich spinels, which have compositions identical to those in less altered komatiites, have survived relatively intact. Shearing effects are developed where the zones are cut by later faults, but over wide areas, primary igneous textures are finely preserved and there is no evidence of shearing, cataclasis or mineral foliation. Overall, these are zones of extreme SiO₂ and K₂O enrichment and MgO and FeO depletion^{5,6} (Table 1).

Crudely stratiform veins, ranging from <1 mm to >5 cm thick, compose as much as 75% of the rock (Fig. 3a). They consist of microcrystalline quartz (chert) and ankeritic carbonate that commonly display a fine cross-vein fibrous structure (Fig. 3a). Most fibres are straight, although not necessarily perpendicular to vein walls, but some are sigmoidal. The veins contain or are separated by chunks of wall rock, and many are bounded above and below or truncated laterally by non-fibrous cavity-filling chert and druse quartz (Fig. 3a). These veins apparently developed by crack-sealing in response to vertical extension^{7,8}.

In many areas, the altered rock contains quartz nodules (Fig. 3b). These include small quartzine (length-slow chalcedony) nodules, generally <2 mm in diameter, and large nodules, up to 5 cm in diameter, consisting of curved, radiating, plumose aggregates of length-slow pseudofibrous or flamboyant quartz.

The textures of the vein-filling chert and carbonate indicate that they are pseudomorphs of earlier fibrous constituents. The following features suggest that one of these constituents was

Table 1 X-ray fluorescence analyses of altered and fresh komatiitic flow rocks, Barberton Greenstone Belt

	1	2	3	4
SiO ₂	85.61	52.38	82.52	52.35
Al ₂ O ₃	8.30	7.16	9.44	10.64
FeO ^T	1.64	12.40	1.95	10.57
MgO	1.07	16.33	1.79	12.79
CaO	0.18	9.37	0.46	11.48
Na ₂ O	0.05	1.11	0.08	1.14
K ₂ O	2.45	0.08	2.98	0.05
TiO ₂	0.53	0.69	0.76	0.59
P ₂ O ₅	0.12	0.08	—	0.05
MnO	0.01	0.23	0.03	0.20
Cr	2,399	1,486	4,387	2,007
Ni	188	289	849	348
Zr	41	48	64	38
Y	21	16	12	14

All Fe expressed as FeO; dry weight normalized to 100%.

1. Altered coarse-spinifex komatiite from uppermost alteration zone shown in Fig. 1. Analysis by J. Veizer and R. Hartree, Department of Geology, University of Ottawa.

2. Fresh coarse-spinifex komatiite from Fig Tree Group. Analysis by G. Byerly, Department of Geochemistry, University of Cape Town.

3. Altered komatiitic basalt from near middle of Kromberg Formation, Komati River. Analysis by J. Veizer and R. Hartree.

4. Fresh basaltic komatiite below sample 3. Analysis by G. Byerly.

gypsum. (1) Within the Onverwacht Group, primary grains of coarse quartz, such as quartz phenocrysts in felsic volcanic rocks and volcanoclastic sandstones and cavity-filling druse quartz, are preserved without alteration. The microcrystalline character of the fibre-replacing quartz suggests replacement of a more poorly ordered silica type, such as chalcedony, or of a non-quartzose mineral, such as carbonate or gypsum. (2) The widespread dissolution of the fibrous veins to form cavities later filled by non-fibrous silica suggests that the original vein material was highly soluble, unlike quartz and ankeritic carbonate. (3) As noted by Ramsay and Huber⁷, 'the length of the fibre [in extensional veins] generally has no special orientation with respect to the crystallographic directions of the crystal making up the fibre.' Virtually all fibre-replacing quartz in the alteration zones, however, shows length-slow aggregate polarization suggesting that it is a pseudomorph of a fibrous length-slow precursor, probably chalcedony. Microcrystalline quartz in the adjacent wall rock is randomly oriented. The associated quartz nodules also consist of length-slow fibrous and pseudofibrous quartz. Whereas most chalcedony is length fast, chalcedony replacing sulphate and evaporites is characteristically length slow⁹. Quartzine spherules and flamboyant quartz are also common silica replacement products of anhydrite nodules^{10,11}. We suggest that stratiform veins in these alteration zones were formed as satinspar or satinspar-calcite crack-sealing veins, similar to fibrous gypsum and carbonate veins widely developed at shallow levels in Phanerozoic sequences¹², and that the quartz nodules represent silicified anhydrite nodules.

These distinctive alteration zones are best developed at the tops of komatiitic flow or pyroclastic sequences that are overlain directly by chert units >1 m thick. Although zones of silicification and K₂O enrichment occur beneath cherts at the tops of tholeiitic and felsic sections, networks of stratiform veins and silica nodules are present only where the host rock is komatiite. Primary host-rock lithology, eruptive style or eruptive setting therefore played a critical role in their formation.

The upper contacts of the alteration zones are generally flat but locally show up to 6 m of erosional relief. Overlying silicified sediments commonly rest with depositional contact on altered komatiite and contain komatiitic clasts and detrital chrome-rich spinels. Beneath depositional contacts, the uppermost 10–50 cm

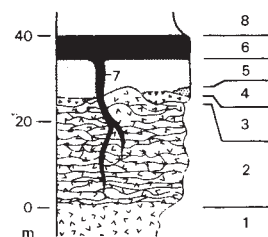


Fig. 2 Idealized sequence of lithologic units associated with alteration zones at the tops of komatiitic flow units. The areal distribution of alteration zones has been outlined previously (Fig. 3b in ref. 2). (1) serpentinized komatiite, (2) alteration zone with crudely stratiform veins, (3) breccia at top of alteration zone, (4) lenticular units of coarse-grained, poorly sorted komatiitic sandstone, (5) silicified volcaniclastic and pyroclastic deposits, (6) black and banded chert, (7) chert dikes, and (8) overlying volcanic or sedimentary units.

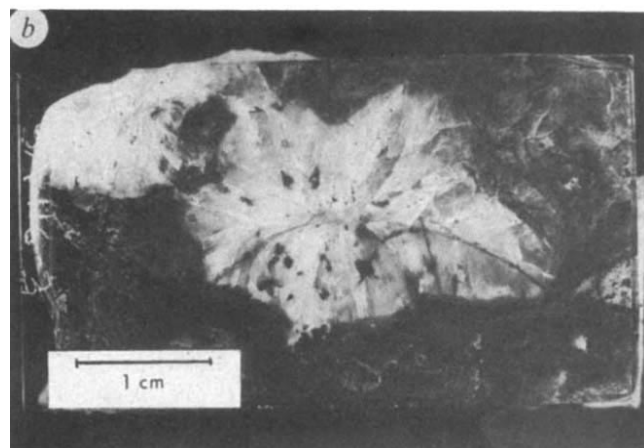
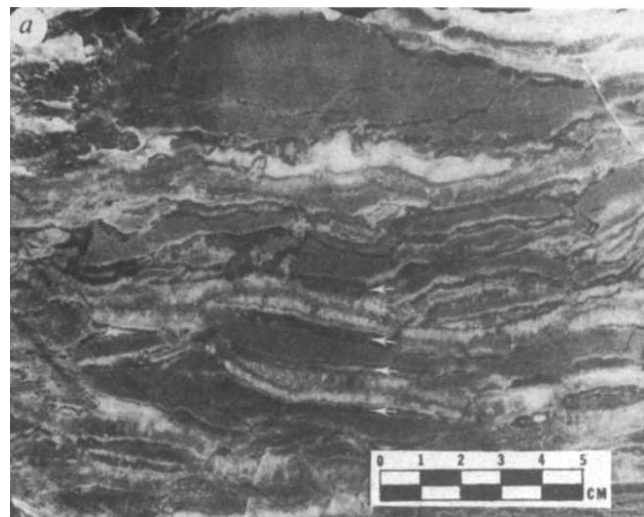


Fig. 3 *a*, Altered komatiitic flow rock showing development of crudely stratiform, fibrous veins. Medium-gray granular masses are altered komatiite remnants. Parallel layers of cavity-filling, non-fibrous chert (dark) and druzy quartz (white) adjacent (arrows) to many veins suggest at least two generations of cracking and vein development. *b*, nodules of plumose flamboyant length-slow quartz in silicified komatiite from alteration zone in the lower part of the Fig Tree Group.

of komatiite are usually brecciated (unit 3, Fig. 2). The breccias consist of angular clasts of unveined komatiite separated by dilatational chert-filled fractures that commonly contain volcaniclastic debris and carbonaceous matter similar to overlying sedimentary units. These breccias are unlike volcanic flow-top breccias and appear to have formed by *in-situ* dilatational fracturing.

Capping cherts (unit 5, Fig. 2) include silicified carbonaceous and volcaniclastic layers deposited mainly under low-energy, shallow-water conditions^{13,14}. Locally, lenses of coarse-grained, poorly sorted, cross-stratified komatiitic debris (unit 4) or stromatolites in growth position rest directly on the komatiites or komatiitic breccias, reflecting more energetic shallow-water to emergent conditions¹⁵. The absence of chunks of stromatolite and sedimentary rock in the breccias indicates that brecciation occurred before or during deposition of the immediately overlying sediments. Many sedimentary layers contain pseudomorphs after crystals of early diagenetic gypsum¹³.

The upper parts of these sedimentary layers generally include units of black chert representing silicified carbonaceous sediments (unit 6, Fig. 2). In several areas, chert dikes (unit 7), formed by the intrastratal flowing of fluid carbonaceous oozes, extend downward from these units, through underlying sedimentary layers, and into the alteration zones. No dikes have been observed extending upward from black chert layers. These dikes truncate veins in the alteration zones, indicating that vein formation preceded dike. Both dikes and veins apparently formed before deposition of more than a few tens of metres of suprajacent sediment.

These alteration zones have been variously interpreted as altered silicic tuffs^{1,16}, cataclastic units marking faults^{3,8}, zones of relatively high-temperature hydrothermal alteration^{5,6}, and weathering zones². Preserved igneous textures and magmatic chrome spinels indicate that the original rocks were komatiitic, not felsic volcanics². Indications that the formation of these zones involved little or no shearing are: the presence of abundant, unstrained spinifex (1), the common absence of mineral foliation and cataclasis (2), their regionally stratiform non-discordant geometry over outcrop distances of up to 25 km (3), their widespread depositional contact with suprajacent sedimentary units (4) and the local presence of earlier-formed vertical veins cut but not offset by the stratiform veins (5) (Fig. 6b of ref. 2).

Our results suggest that these alteration zones formed through the following sequence of partly overlapping events. (1) The rapid eruption of thick sequences of komatiitic lavas in overall shallow-water settings resulted in a widespread build-up of the volcanic surfaces into the subaerial environment. (2) Post-eruption subaerial exposure of the flow tops was accompanied by erosion and the local accumulation of komatiitic alluvial sand

and gravel (unit 4, Fig. 2). This may also have been a time of chemical weathering of the flow tops. (3) Regional subsidence in the absence of volcanism led eventually to submergence. Passage of the flow tops through marginal evaporitic environments was associated with planation, formation of anhydrite nodules, brecciation, and local growth of stromatolites. (4) Shallow-water deposition of volcaniclastic and carbonaceous sediments resulted in infiltration of the breccias by loose debris and the accumulation of sediment caps on the komatiitic flows. (5) Silicification of fine-grained volcaniclastic units began almost concurrently with deposition^{13,17}, forming nearly impermeable chert caps above the alteration zones. Silicification of the upper parts of the lavas may also have begun at this time. Fluids moving through the volcanic rocks were trapped beneath these seals resulting in hydraulic overpressuring and formation of crack-seal veins. These were low-temperature fluids, probably less than 60 °C, as suggested by the precipitation of gypsum¹⁸. They may have been dense sulphate-rich brines originating in marginal evaporitic environments. (6) The formation of chert dikes reflects fracturing of the chert caps above the alteration zones and may mark the release of hydraulic overpressures and the end of vein formation. Later changes in pore-water chemistry, perhaps related to burial or the upward movement

of deeper-level, higher-temperature fluids, resulted in gypsum replacement and later stages of silicification, carbonatization and K_2O enrichment. Where the chert caps remained intact, the formation of additional crack-seal veins may have accompanied these metasomatic events.

The results of this study indicate that these distinctive Archaean alteration zones formed initially through a complex sequence of surface and near-surface low-temperature processes. Later metasomatism and metamorphism also played a major role in their final character. Studies of this and other flow-top alteration zones may provide important clues to Archaean water and atmospheric chemistry, surface and subsurface hydrologic systems, and the relationships among volcanism, subsidence and sedimentation in greenstone belts.

This research was supported by NSF grants EAR 79-19908, EAR 82-06015 and EAR 84-06420 to D.R.L. and EAR 79-19909 to G.R.B. and D.R.L. We thank the Department of Geology, University of the Witwatersrand, and the Anglo American Corporation of South Africa, for field support.

Received 27 March; accepted 29 July 1986.

1. Viljoen, M. J. & Vilgoen, R. P. *Spec. Publ. geol. Soc. S. Afr.* **2**, 113-152 (1969).
2. Lowe, D. R., Byerly, G. R., Ransom, B. L. & Nocita, B. W. *Precamb. Res.* **27**, 165-186 (1985).
3. de Wit, M. L. *J. struct. Geol.* **4**, 117-136 (1982).
4. de Wit, M. J., Hart, R., Martin, A. & Abbott, P. *Econ. Geol.* **77**, 1783-1801 (1982).
5. Pearton, T. N. in *Komatitites* (eds Arndt, N. E. & Nisbet, E. G.) 459-476 (Allen & Unwin, London, 1982).
6. Hanor, J. S. & Duchac, K. C. *Abstr. Prog. geol. Soc. Am.* **17**, 603.
7. Ramsay, J. G. & Huber, M. I. *The Techniques of Modern Structural Geology Vol. 1: Strain Analysis*, 235-263 (Academic, London, 1983).
8. de Wit, M. J. in *Abstr. Workshop tecton. Evol. Greenstone Belts* **25**, (Lunar & Planetary Institute, Houston, 16-18 Jan. 1986).
9. Folk, R. L. & Pittman, J. S. *J. sedim. Petrol.* **41**, 1045-1058 (1971).
10. Chowins, T. M. & Elkins, J. E. *J. sedim. Petrol.* **44**, 885-903 (1974).
11. Tucker, M. E. *Sedim. Geol.* **16**, 193-204 (1976).
12. Machel, H. *Sedimentology* **32**, 443-454 (1985).
13. Lowe, D. R. & Knauth, L. P. *J. Geol.* **85**, 699-723 (1977).
14. Lanier, W. P. & Lowe, D. R. *Precamb. Res.* **18**, 237-260 (1982).
15. Byerly, G. R., Lowe, D. R. & Walsh, M. M. *Nature* **319**, 489-491 (1986).
16. Anhaeusser, C. R. *Phil. Trans. R. Soc. A273*, 359-388 (1973).
17. Paris, I., Stanistreet, I. G. & Hughes, M. J. *J. Geol.* **93**, 111-129 (1985).
18. Hardie, L. A. *Am. Miner.* **52**, 171-200 (1967).

Body size, ecological dominance and Cope's rule

James H. Brown & Brian A. Maurer*

Department of Ecology and Evolutionary Biology,
University of Arizona, Tucson, Arizona 85721, USA

We present data and analyses demonstrating that large species utilize a disproportionately large share of the resources within local ecosystems. Even though small species tend to have higher local population densities, these are not sufficient to compensate for their lower rates of energy use per individual. The relationship is very general; holding for example for birds, mammals, fish and plants. We suggest that several ecological advantages enable larger species and larger individuals within species to monopolize resources, and that the resulting selection pressures are responsible for the evolutionary trend towards increasing body size seen in many phyletic lineages. Our results contradict important studies¹⁻⁴ that have concluded that species of small body size use at least as large a proportion of the resources within ecosystems as their larger relatives.

A goal of community ecology is to understand the patterns and processes that determine the numbers and variety of coexisting species. Are there general rules that characterize the contributions of the different species to the structure and dynamics

of ecosystems? One measure of the ecological importance of species is their relative utilization of energy and transfer of materials^{5,6}. Within groups of taxonomically-related, ecologically similar organisms, two variables (individual body size and population density) are of primary importance in determining how energy and other limiting resources are allocated. Large individuals demand more from environmental resource pools than small ones, but also tend to occur at lower population densities. The question, then, is to what extent the greater abundance of small species compensates for their lower resource requirements per individual.

Although direct measurements of energy flow per unit area are probably not available for all species in any community, these rates can be estimated quite accurately from data on population density (D), individual body mass (M) in animals and photosynthetic cover (C) in plants. In animals the average daily energy use of a free-living individual (U) scales allometrically as approximately $M^{0.67}$ (refs 7, 10), so total energy use per species (E) is given as $E = DU = DM^{0.67}$. In plants, variation in photosynthetic rate per unit leaf surface among species is insignificant compared to the total leaf surface^{11,12}, so the energy use of each species is directly proportional to its vegetative cover (C).

We analysed data on D and M for local groups of birds, mammals and fish, and on C for communities of plants to determine how rates of energy use per species (E) varies with body size. Using the results of 39 censuses of breeding land birds and data on species body mass¹³, we tried three different methods¹⁴ to fit a linear relationship for $\log D$ with $\log M$ (Table 1). Depending on the method, only 3 to 10 of the 39 slopes were ≤ -0.67 . Pooling the data gave a slope of approximately -0.30 , not greatly different from the value of -0.19 reported by Peters². If D scales as $M^{-0.30}$ and U scales as $M^{0.67}$, then the rate of energy intake per species ($E = DU$) increases with body mass, scaling as $\sim M^{0.37}$. Because the three methods of estimating slopes give comparable results, in the following analyses we present only the results of the regression method applied to log-transformed data.

For the nine species of granivorous rodents (mammals) in a Chihuahuan Desert ecosystem¹⁵, regression of $\log D$ on $\log M$ gave a significant positive relationship ($r = 0.56$, slope [b] = 1.17, $n = 9$, $P = 0.05$). Consequently, in these mammals the rate of energy flow per species ($E = DU$) scales positively with body mass as $\sim M^{1.17} \times M^{0.67} = M^{1.84}$.

We analysed three independent data sets¹⁶⁻¹⁸ on population densities and body masses of marine fish (Table 2). The slopes of regressions of $\log D$ on $\log M$ averaged near zero, and only one was more negative than -0.1 . Thus, if D scales as $\sim M^{0.0}$ and U scales as $M^{0.67}$, then the average rate of energy use per species ($E = DU$) must vary as $\sim M^{0.67}$.

We obtained complete census data for perennial plants in five different habitats in two desert regions^{19,20}. Total cover (C) of each species was used to estimate its contribution to energy flow, and cover divided by population density (D) was taken as the effective size of an individual plant. When we regressed $\log C$ against $\log (C/D)$, all the slopes were positive, with values ranging from 0.53 to 1.59 and averaging ~ 1.0 (Table 3). Although the usefulness of estimating individual mass in perennial plants containing nonliving woody tissue is questionable, M should vary with individual surface area as $\sim (C/D)^{1.5}$ and rate of energy use per species (E) should scale as $\sim M^{1.0}$.

Our analyses show that in four different taxonomic groups (birds, mammals, fish and plants) species comprising large individuals account for most of the energy flow and resource use within local ecosystems. These results contradict those of two influential recent studies. Damuth¹ claimed that population densities of mammals scale as $M^{-0.75}$, so that rates of energy use per species are essentially independent of body size (E scales as $\sim M^{0.0}$). Peters²⁻⁴ compiled data which suggest that population density (D) often scales as $M^{-1.0}$ so that energy use per

* Present address: Department of Zoology, Brigham Young University, Provo, Utah 84602, USA.

Table 1 Slopes of log population density and log body mass¹³ for land birds in different habitats

Habitat	log data regression	log normal data regression	RMA	Correlation coefficient	No. of species
Savanna	-0.53	-0.64	-1.11	-0.40	17
Grassland	0.19	-0.18	1.36	0.45	19
Deciduous forest	1.08	2.47	7.29	0.48	15
Managed forest	-0.05	-0.25	-1.40	-0.04	35
Douglas fir	-0.90	-1.03	-1.93	-0.57	21
Clearcut forest	-0.31	-0.99	-2.86	-0.26	15
Ponderosa pine	-0.18	-0.18	-0.32	-0.26	21
Logged ponderosa	-0.44	-0.57	-0.60	-0.47	22
	-0.31	-0.33	-0.44	-0.39	23
	-0.04	-0.07	-0.10	-0.06	19
Clearcut forest	-0.72	-0.18	-0.12	-0.50	6
Spruce	-0.41	-0.20	-0.20	-0.25	16
	-0.51	-0.17	-0.09	-0.34	21
Spruce/jackpine	-0.58	-0.31	-0.37	-0.32	22
Spruce forest edge	-0.60	-0.32	-0.21	-0.35	12
Spruce/jackpine edge	-0.59	-0.35	-0.44	-0.31	21
Spruce riparian	-0.50	-0.46	-0.40	-0.24	18
Beech/maple	-0.37	-1.34	-1.62	-0.41	29
	-0.37	-1.51	-1.23	-0.51	24
Urban woodlot	0.39	0.73	5.15	0.36	20
	0.17	1.73	-0.28	0.20	17
Pine/oak seral	-0.18	-0.15	-0.51	-0.09	23
Pine/oak mature	-0.67	-0.66	-2.32	-0.31	23
Cove forest	-0.17	-0.14	-0.35	-0.11	31
Hemlock/deciduous	-0.63	-0.33	-0.28	-0.45	26
Chestnut oak	-0.36	-0.13	-0.13	-0.23	32
Red oak	-0.31	-0.10	-0.09	-0.29	25
Pine heathland	-0.57	-0.13	-0.09	-0.39	14
Gray beech	-0.08	-0.17	-0.20	-0.06	13
Spruce/fir	-1.05	-5.68	-5.14	-0.35	11
	-0.34	-4.37	-4.34	-0.25	10
	0.10	-0.24	0.34	0.07	14
	-0.37	-0.60	-0.54	-0.23	15
Grassland	-0.01	-0.13	-0.26	-0.01	13
Brushland	-0.49	-0.27	-0.24	-0.40	14
Savanna	0.02	-0.11	0.56	0.01	14
Shrubland	0.03	-0.31	1.22	0.02	15
Aspen	-0.11	-0.02	-0.01	-0.13	28
Savanna	-0.17	-0.44	-1.34	-0.11	12
Pooled data	-0.30	-0.12	-0.40	-0.27	197

Three methods of calculation were used: linear regression on log-transformed data, linear regression and reduced major axis (RMA) fitted to log-normally distributed data¹⁴. Note that only 3–10 slopes were < -0.67 .

These 39 censuses were taken from the published literature; the references are available from the authors on request.

Table 2 Regression analyses of log population density on log body mass for fish

Habitat	Slope	Standard error	Intercept	Correlation coefficient	No. of species
Caribbean coral reef ¹⁶	-0.05	0.08	0.76	-0.08	53
Gulf of California tide pool ¹⁷	0.04	0.22	2.26	0.04	25
Newport Bay, California ¹⁵					
bag seine	-0.08	0.36	1.18	-0.05	27
seine/bay	0.32	0.65	1.22	0.12	19
seine/shallows	-0.07	0.98	1.96	-0.03	8
drop net	-0.48	0.35	0.89	-0.37	14
enclosure	0.24	0.46	1.11	0.25	6
otter trawl	0.16	0.12	0.60	0.22	35
gill net	-0.03	0.26	0.88	-0.03	22

Note that none of the slopes are more negative than -0.50 .

species should scale negatively with body mass (as $M^{-0.33}$). The results of Damuth and Peters can probably be attributed to a sampling bias inherent in their data collection procedures. Their analyses are based on published values of population densities for species from different, scattered localities. These data probably neglect the numerous rare species of small body size that occur within local ecosystems, because ecologists tend to study

conspicuous and abundant species. We have been careful to avoid this and other possible biases by using only complete censuses of local biotas.

We have calculated the slopes of allometric equations to reveal differences between our results and those of Damuth and Peters. However, the low correlation coefficients of our regressions suggest that no allometric equation adequately characterizes the

Table 3 Regression analyses of log cover on log (cover/population density) for perennial plants in five different habitats in two desert regions of southwestern North America

Habitat	Slope	Standard error	Intercept	Correlation coefficient	No. of species
Sonoran Desert ¹⁹					
Goldeneye	0.55	0.18	0.91	0.53	27
Agave	0.53	0.14	0.67	0.56	33
Mojave Desert ²⁰					
Shadscale	1.32	0.26	3.40	0.76	21
Bladder-sage	1.25	0.17	3.19	0.87	19
Spiny menodora	1.59	0.13	4.36	0.93	23

Note that all the slopes are greater than 0.50.

relationship between population density and body size. Our qualitative result that on average small species use proportionally less energy than large ones can be corroborated using a robust nonparametric technique. We divide the species in a census into those above and below the median body mass for species in that census. Then if population energy use (E) is on average greater for larger than for smaller species, then $\bar{E}_s < \bar{E}_l$, where $\bar{E}$ represents mean energy use of the smaller (s) and larger (l) species. For animals in which individual energy use scales as $M^{0.67}$, we can substitute $\bar{D}\bar{M}^{0.67}$ for $\bar{E}$, so $\bar{D}_s\bar{M}_s^{0.67} < \bar{D}_l\bar{M}_l^{0.67}$. Because the probability that any one census obeys this relationship by chance is 0.5, whenever we have a sufficient number of censuses we can use a binomial test to evaluate the hypothesis that the collection of censuses exhibits the expected pattern. For the bird censuses in Table 1, 35 out of 39 showed the predicted relationship (one-tailed test, $p < 0.0001$).

We have shown that on a per species basis, within trophic levels and taxonomic groups, organisms of large individual size account for most of the energy flow through local ecosystems. This has important ecological and evolutionary implications. There are several ecological advantages of large body size. Large species are usually dominant in interspecific aggression²¹, which may result in exclusion of small species from preferred resources; they tend to have more efficient homeostatic mechanisms and greater mobility^{2,8,10}, so they are able to tolerate a wider range of environmental conditions and to seek out more favourable locations. An increase in size enables an individual to spend less energy per unit biomass on maintenance and to become more efficient at extracting usable energy from low-quality foods^{2,22}. Consequently, the same amount of available energy can support a greater biomass of a large species than of a small one. Finally, size may be important in avoiding predators²³. Our data suggest that these advantages enable large species to dominate the resource allocation in ecosystems, leaving smaller species to divide up the remainder.

The ecological consequences of body size provide a mechanistic explanation for the evolutionary cycle known as Cope's rule^{23,24}. Phylogenetic lineages usually begin with relatively small organisms that give rise to larger forms. Ultimately giant forms are produced, but these frequently become extinct, creating new opportunities for another lineage as it in turn evolves organisms of larger size. Presumably the ecological advantage of monopolizing resources provides the selective pressure that promotes evolution of greater size. Individuals of large size are favoured by intraspecific natural selection, because they can dominate resource use and consequently leave more offspring than their smaller relatives. However, this increase in body size is accompanied by a higher probability of extinction, which may be attributed to lower population density, smaller population size, slower population growth and perhaps other factors^{23,24}. Thus Cope's rule is the result of two opposing evolutionary processes, one operating at the level of individuals within populations, and the other at the level of species within ecosystems²⁵.

We thank M. Kurzius and R. Robichaux for technical assistance; W. A. Calder III, J. Damuth, P. Harvey, C. Istock, D.

Jablonski, L. Van Valen and an anonymous reviewer for valuable comments; the University of Arizona for computer funds; and the NSF for research support grant BSR 8021535 to J.H.B.

Received 7 October 1985; accepted 26 September 1986.

1. Damuth, J. *Nature* **290**, 699-700 (1981).
2. Peters, R. H. *The Ecological Implications of Body Size* (Cambridge University Press, Cambridge, 1983).
3. Peters, R. H. & Wassenberg, K. *Oecologia* **60**, 89-90 (1983).
4. Peters, R. H. & Raelson, J. V. *Am. Nat.* **124**, 498-517 (1984).
5. Odum, E. P. *Am. Zool.* **8**, 11-18 (1968).
6. Odum, H. T. *Systems Ecology* (Wiley, New York, 1983).
7. Garland, T. *Am. Nat.* **121**, 571-587 (1983).
8. Mace, G. M. & Harvey, P. H. *Am. Nat.* **121**, 120-132 (1983).
9. Walsberg, G. E. in *Avian Biology* vol. 6 (eds Farner, D. S. & King, J. R.) 161-220 (Academic, New York, 1983).
10. Calder, W. A. *Size, Function, and Life History* (Harvard University Press, Cambridge, 1984).
11. Mooney, H. A. *A. Rev. ecol. Syst.* **3**, 315-346 (1972).
12. Schulze, E. D. in *The Encyclopedia of Plant Physiology*, new series, vol. 126, 616-676 (Springer, Berlin, 1982).
13. Dunning, J. B. *Western Bird Banding Assoc. Monogr.* **1**, 1-38 (Eldon, Cave Creek, 1984).
14. Kermack, K. A. & Haldane, J. B. S. *Biometrika* **37**, 30-41 (1950).
15. Brown, J. H. & Munger, J. C. *Ecology* **66**, 1545-1563 (1985).
16. Smith, C. L. & Taylor, J. C. in *Results of the Tektite Program: Ecology of Coral Reef Fishes* vol. 14 (eds Collette, B. B. & Earl, S. A.) 125-170 (Science Bulletin, Natural History Museum of Los Angeles County, 1972).
17. Thomson, D. A. & Lehner, C. E. *J. exp. mar. biol. Ecol.* **22**, 1-28 (1976).
18. Horn, M. H. & Allen, L. G. *California Department of Fish and Game Tech. Rep.* **45**, 1-36 (1981).
19. Turner, R. M. & Brown, D. E. in *Desert Plants* vol. 4 (ed. Brown, D. E.) 181-221 (University of Arizona, Tucson, 1982).
20. Kurzius, M. A. Thesis, Univ. Nevada, Las Vegas, 1981.
21. Morse, D. H. *Behavioral Mechanism in Ecology* (Harvard University Press, Cambridge, 1980).
22. Sibley, R. M. in *Physiological Ecology* (eds Townsend, C. R. & Callow, P.) 109-139 (Blackwell, London, 1981).
23. Stanley, S. M. *Macroevolution* (Freeman, San Francisco, 1979).
24. Bakker, R. T. in *Patterns of Evolution* (ed. Hallam, A.) 431-478 (Elsevier, Amsterdam, 1977).
25. Fowler, C. W. & MacMahon, J. A. *Am. Nat.* **119**, 480-498 (1982).

Mach bands are phase dependent

M. Concetta Morrone, John Ross, David C. Burr & Robyn Owens*

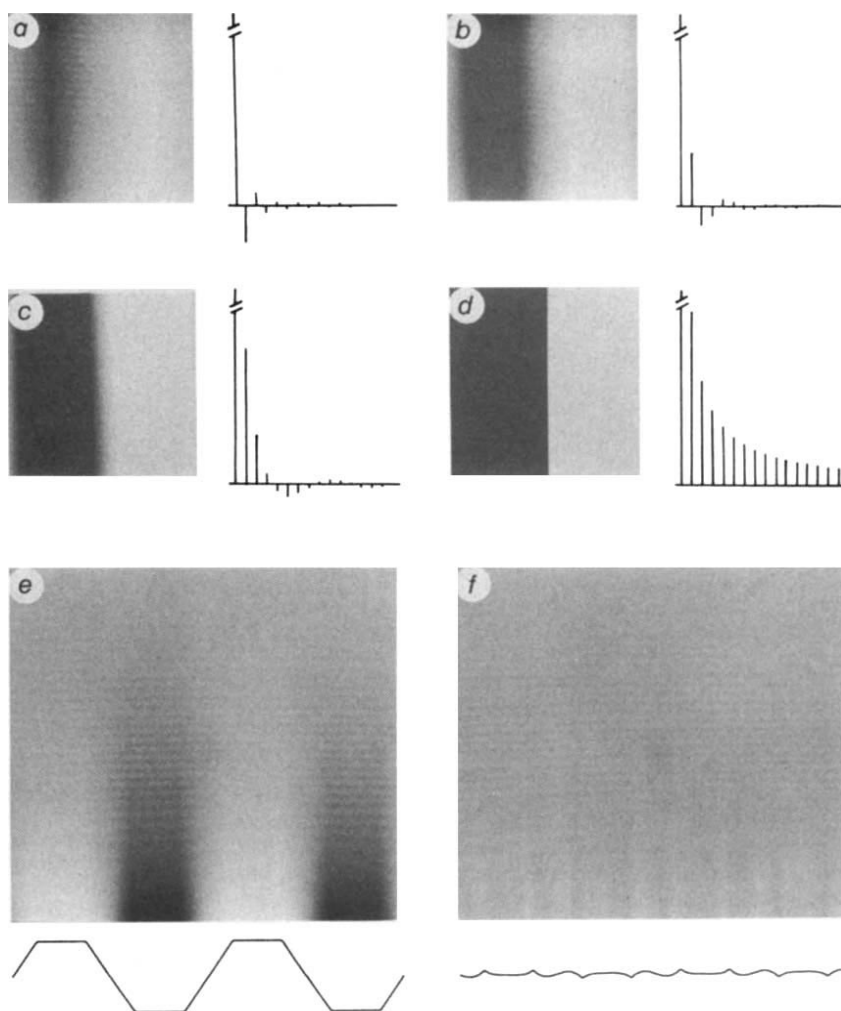
Department of Psychology and * Department of Computer Science, University of Western Australia, Nedlands, Western Australia 6009, Australia

The prevailing explanation of Mach bands, the paradoxical bands of light and dark seen where luminance gradients meet plateaux, is that they are due to lateral inhibition in the visual system¹⁻⁴. This explanation equates Mach bands with distortions in a processed luminance distribution due to selective attenuation of low frequency components. But square waveforms exhibit no Mach bands^{2,5,6}, although they should also be distorted after processing. Measurements of the contrast required to see Mach bands in trapezoidal waveforms and manipulations of their spectra lead us to conclude that phase relationships between Fourier components are important to the structure we perceive. A model based on the odd and even symmetry of visual receptive fields explains our results.

Fig. 1 *a-d*, Examples of four types of trapezoids. The shape of the waveform is determined by the parameter t (the ratio of the ramp width to the period), which is 0.5 in *a* (triangular wave), 0.25 in *b*, 0.125 in *c*, and 0 in *d* (square wave). Mach bands are clearly visible in all except the square wave. The Fourier expansions of the patterns are given by

$$F(x) = \sum_{k=0,\infty} [(4A/t\pi^2) \cdot (\sin(\pi t(2k+1)) / (2k+1)^2)] \cdot \sin(2\pi(2k+1)x/T), \quad (1)$$

where A is the amplitude and T is the period. The spectra are depicted schematically adjacent to each waveform (the first harmonic terms are not to scale). Only the odd terms are non-zero, and they fall into alternatively positively and negatively weighted blocks whose size varies inversely with t . *e, f*, At left a trapezoid waveform ($t=0.25$), decreasing in contrast exponentially upwards. At right the residual waveform, from which the first block of positively weighted harmonics (the first and the third) has been removed, also decreasing in contrast exponentially upwards. For various distances (spatial frequencies), note the point at which the Mach bands disappear from the trapezoid, and also the point at which the residual waveform falls below detection threshold. The two thresholds should be similar. By fixating some distance below the two panels that Mach bands disappear in the periphery as contrast decreases well before the residual ceases to be visible.



Mach originally⁷ used rotating discs and drums to study the bright and dark bands that appear where ramps of luminance meet plateaux as, for example, at the borders of the penumbra of a shadow. He later devised optical methods to cast shadows⁸. For the studies reported here we used gratings formed by computer on an oscilloscope which were variable in luminance profile and contrast. Fig. 1 shows four trapezoids, including the two extreme cases of a square wave ($t=0$) and the triangular wave ($t=0.5$), where t is the ratio of the width of each ramp to the period. Classical Mach bands are seen on the trapezoids (Fig. 1*b, 1c*); bright bands are visible at the top of each ramp and dark bands at the bottom. Note also that the ramps do not appear to have a uniform brightness gradient and the plateaux do not appear uniformly dark and light. The triangular wave also exhibits sharp bands at the peaks and troughs of the waveform, but the square wave does not.

The Fourier transforms of the trapezoid waveforms (calculated from the point where the waveform crosses zero) are also shown in Fig. 1. The components are arranged in blocks alternatively positively and negatively weighted (that is, differing in phase by 180°). The number of components in each block varies inversely with t (see formula in the legend), being one for the triangular wave, and infinite for the square wave. The observations that Mach bands are not seen on square waves^{2,6}, which have no out-of-phase components, or on trapezoids with very steep ramps⁵, where the out-of-phase components are all of high spatial frequency, led us to suspect that phase may hold the key to Mach bands.

We isolated residual waveforms from trapezoids by removing the first block of positive harmonics (see Fig. 1*f*) and measured

the contrast necessary to see them as a function of spatial frequency (inverse of their period, T in equation (1), Fig. 1, legend). The results are shown by the filled circles in Fig. 2. We then measured the minimum contrast at which Mach bands could be seen in the original trapezoidal waveforms (open triangles in Fig. 2). The two curves follow one another closely as spatial frequency is varied. Interestingly, observers spontaneously remarked that the two tasks, setting thresholds for seeing Mach bands and for detecting residual waveforms, seemed to make common demands on them, as if the two tasks were almost identical. Readers can confirm some of the results for themselves by observing the contrast at which the Mach bands of Fig. 1*e* and the pattern of Fig. 1*f* disappear at various viewing distances.

This result suggests that Mach bands are visible only when the out-of-phase Fourier components (together with all higher harmonics) reach their independent threshold. To test this more directly, we attenuated the higher frequency components by applying a low-pass gaussian filter to various trapezoidal waveforms and decreased the space constant of the gaussian filter until the Mach bands disappeared. We then applied the same filter to the residual waveforms. For all the trapezoids we investigated, the low-pass filter which removed the Mach bands attenuated the appropriate residual waveform to a contrast just below its detection threshold.

It is known that phase discrimination deteriorates in the periphery^{9,10}. We measured the contrast required to see Mach bands in various trapezoids and the contrast required to see their residuals as a function of retinal eccentricity. As eccentricity increased the contrast required to see Mach bands increased

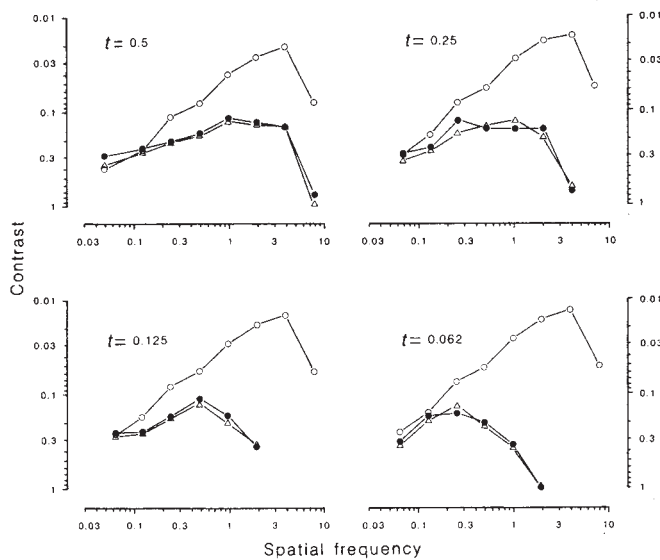


Fig. 2 Contrast required to see Mach bands as a function of spatial frequency (c per deg) is shown by open triangles, for four different t values. Contrast required to see the residual waveform (first block of positive components removed) is shown by the filled circles. The curves follow one another closely. Open circles show contrast required to see the fundamental of the trapezoid, for comparison. All contrasts are expressed as those of trapezoids from which the waveforms derive.

much more rapidly than that required to see the residual. Indeed, Mach bands could not be seen at all at eccentricities greater than three degrees (as previously observed by Mach⁸), whereas the residual was still clearly visible. Phase sensitivity is indeed poor in the periphery; when viewed peripherally, trapezoids become unstable and alternate in appearance between square and triangular waves.

As a final confirmation of the importance of phase, we created synthetic waveforms by inverting the sign of all negatively weighted components of the triangular wave ($t=0.5$) and a trapezoid ($t=0.25$), to make 'all positive' waveforms, having the same amplitude spectra as the parent trapezoids (Fig. 3). As can be seen, the luminance profiles are now smooth. There are no Mach bands in either case, but sharp edges are apparent, similar to those of a square wave.

The effects of phase on the appearance of distributions of luminance are more radical when there is variation in two dimensions. The distribution of luminance in Fig. 3c is pyramidal. In both horizontal and vertical cross-section, the luminance profile is triangular. The Mach bands of one dimensional triangular waveforms reappear in two dimensions as star-like patterns where the two dimensional pattern forms points of high and low luminance. Adjacent is the same pattern with all phases set positive. The structure is now like that of a checkerboard, with white and black squares clearly separated by sharp vertical and horizontal borders.

The results of this study suggest that Mach bands are different from edges, but that both depend on phase relationships. When all negative harmonics (180° out-of-phase) are rendered invisible (by lowering their contrast, increasing their spatial frequency, or smooth digital filtering), Mach bands disappear. When patterns are displayed peripherally (where the visual system is poor at resolving phase^{9,10}) the bands also disappear. Finally, if all negative harmonics of a trapezoid are shifted in phase by 180° , Mach bands disappear, but an edge is seen where there was none before.

As mentioned earlier, all harmonics of a square wave are in zero phase at the point where the waveform crosses zero in a

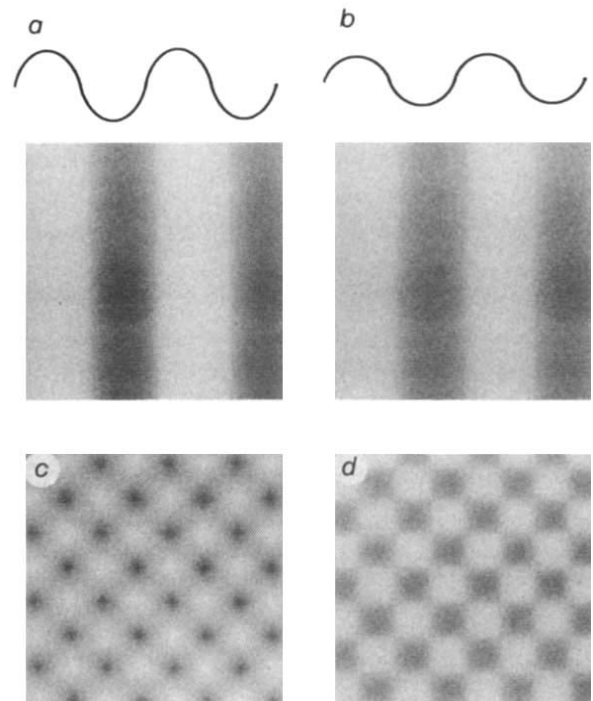


Fig. 3 *a, b*, Synthetic waveforms constructed by inverting the sign of all negatively weighted components of trapezoids of t value 0.5 (*a*) and 0.25 (*b*). The waveforms have no discontinuities, yet we perceive sharp edges, with no Mach bands; *c*, two dimensional pattern formed by multiplying a horizontal by a vertical triangular wave, shows white and black star patterns which are two-dimensional analogues of Mach bands; *d*, two dimensional pattern derived from *c* by inverting the phase of all negatively weighted harmonics. The star patterns reflecting two-dimensional Mach bands are no longer visible. Instead, a checkerboard-like pattern is seen, with sharp borders between the black and white squares.

positive direction and 180° where it crosses zero in a negative direction. This is also true for 'all positive' trapezoids in which all out-of-phase components have been shifted in phase by 180° . Elsewhere, the phases of the harmonics advance at different rates, so no other points have strong congruence of phase. Trapezoidal waveforms other than the square wave have blocks of negatively weighted harmonics which break the congruence of phase at the zero-cross point. However, the periodic halt of the phase advance of the harmonics causes the phases to pile up around $\pm 90^\circ$ at the points where the ramps meet plateaux. A phase grouping around 90° is typical of that produced by delta functions and bars and may be the signal that produces Mach bands.

Why should the visual system react so strongly to phase relationships? A two dimensional basis is required to extract phase information. The even and odd symmetry of receptive fields of the visual cortex¹¹ could form such a basis. The receptive fields vary in size, also allowing for coding of an image at different spatial scales¹²⁻¹⁵. The potential ability of such a system to signal edges (odd-symmetric) and lines (even symmetric) has been noted by many investigators¹⁵⁻¹⁸. We suggest that an efficient means for the visual system to locate bars and edges would be to consider the sum of the squared output of even and odd symmetric filters, which always peaks at points of phase congruence. After the peak had been located the response of the odd and even fields at that point would determine if it were due to an edge or a bar, and give the sign and contrast of the edge or bar. The clustering of phase around 90° , which occurs with trapezoids, should also signal a bar or stripe.

Phase is thought to be encoded by even and odd symmetric detectors in the human visual system¹⁹. It has further been shown that the sensitivity of the odd, but not the even, symmetric mechanism is reduced in peripheral vision²⁰. This would account for the disappearance of Mach bands in the periphery, at contrasts well above those required for the independent detection of out-of-phase harmonics.

Although the generally accepted explanation of Mach bands is that they are due to lateral inhibition, several investigators have noted the problem this presents for square waves and sharp edges^{2,5,6,16} and have considered other possibilities. Tolhurst observed that even symmetric receptive fields (which he termed bar detectors) may signal Mach bands¹⁶. He also speculated that for the square wave the response of odd symmetric fields (edge detectors) may inhibit that of the even symmetric fields (bar detectors), so no stripes are seen (see also ref. 6). Watt and Morgan's²¹ general theory of spatial vision also predicts bar signals at the border of a ramp, provided that they are far enough apart. Both these ideas have some similarity to ours, but our results (and simulations to be reported in a fuller paper) suggest that inhibition between edge detectors and bar detectors is unnecessary.

Finally, we note that trapezoids, like square waves and synthetic 'all positive' trapezoids, show zero-crossings at all scales at the mean luminance point. If edges were encoded by align-

ments of zero-crossings, as has been assumed²²⁻²⁴, all should have edges at that point. But trapezoids do not. A model seeking phase congruence as the signature of bars and edges, built on adequate basis functions, can without ambiguity, locate edges and bars where we see them.

Received 2 June; accepted 2 October 1986.

1. Ratliff, F. *Mach Bands* (Holden-Day, San Francisco, 1965).
2. von Békésy, G. *Vis. Res.* **8**, 1483-1499 (1968).
3. Cornsweet, T. N. *Visual Perception* (Academic, New York, 1970).
4. Fiorentini, A. in *Handbook of Sensory Physiology* VII/4, 188-201 (Springer, Berlin, 1972).
5. Ross, J., Holt, J. J. & Johnstone, J. R. *Vis. Res.* **21**, 1165-1167 (1981).
6. Ratliff, F. *Vis. Res.* **24**, 163-166 (1984).
7. Mach, E. *I.S.-B. Akad. Wiss. Wien. math.-nat. Kl* **54**, 303-322 (1865) (translated in ref. 1).
8. Mach, E. *I.S.-B. Akad. Wiss. Wien. math.-nat. Kl* **115**, 633-648 (1906) (translated in ref. 1).
9. Rentschler, I. & Treutwein, B. *Nature* **313**, 308-310 (1985).
10. Braddick, O. J. *Doc. Ophthal. Proc. Series* **30**, 255-262 (1981).
11. Hubel, D. H. & Wiesel, T. N. *J. Physiol.* **160**, 106-154 (1962).
12. Maffei, L. & Fiorentini, A. *Vis. Res.* **17**, 257-264 (1977).
13. De Valois, R. L., Albrecht, D. G. & Thorell, L. G. *Vis. Res.* **22**, 545-559 (1982).
14. Maffei, L., Morrone, M. C., Pirchio, M. & Sandini, G. *J. Physiology* **296**, 27-47 (1979).
15. Kulikowski, J. I. & Bishop, P. O. *Expl. Brain Res.* **44**, 386-400 (1981).
16. Tolhurst, D. J. *Vis. Res.* **12**, 797-804 (1972).
17. Pollen, D. A. & Ronner, S. F. *Science* **212**, 1409-1411 (1981).
18. Pollen, D. A. & Ronner, S. F. *Vis. Res.* **22**, 101-118 (1982).
19. Field, D. J. & Nachmias, J. *Vis. Res.* **24**, 333-340 (1984).
20. Bennett, P. & Banks, M. *Nature* (submitted).
21. Watt, R. J. & Morgan, M. J. *Vis. Res.* **24**, 1387-1397 (1984).
22. Marr, D. *Vision* (Freeman, San Francisco, 1980).
23. Marr, D. & Hildreth, E. *Proc. R. Soc. Lond. B* **207**, 187-217 (1980).
24. Hildreth, E. C. *MIT. A.I. Lab. Memo* **858**, (1985).

Hysteresis in the perception of motion direction as evidence for neural cooperativity

Douglas Williams, Gregory Phillips & Robert Sekuler

Departments of Neurobiology/Physiology and Psychology,
Northwestern University, Evanston, Illinois 60201, USA

When elements of a parallel network, such as the human brain, are extensively interconnected, the network can exhibit 'cooperative behaviour'. Such behaviour, which is characterized by order-disorder transitions, multi-stable states, and a form of memory called 'hysteresis', has been observed in human stereopsis^{1,2} and has motivated models of stereopsis that incorporate cooperative networks³⁻⁶. More recently, cooperative phenomena have also been observed in human visual motion perception⁷. This report strongly supports a cooperative interpretation of motion perception by demonstrating hysteresis in the perception of motion direction. The results agree quantitatively with a mathematical model incorporating nonlinear excitatory and inhibitory interactions among direction-selective elements.

Our stimuli were dynamic random dot cinematograms comprising 512 computer-generated dots. Each dot took an independent, two-dimensional random walk of constant step size (0.9°). The direction in which any dot moved was independent of its own previous displacements and also of the displacements of the other dots⁸.

The direction of motion for each dot was chosen from either (1) one of two uniform distributions, or (2) a mixture of these two distributions. When all dots drew their movements from a uniform distribution extending over 360°, only the 'local random motion' of individual dots was evident. This resembled a detuned television set. However, when all dots drew their movements from a uniform distribution of 180° or less, the dots appeared to flow *en masse* in the direction of the distribution mean—although individual perturbations were still evident. This resembled snowflakes that, although individually perturbed by wind currents, seem to drift in one direction. In all experiments the mean of the signal distribution was upward.

For convenience, we refer to the 180° and 360° distributions as the signal and noise distributions, respectively. When dots drew their movements from a combination of these distributions, each dot's displacement came randomly from either distribution. The resulting perception depended upon the relative proportions of noise and signal.

Two types of trials were run in random order. In one, the direction of motion for each dot was chosen initially from the signal distribution, producing a perception of upward flow. After a random interval lasting up to 12 seconds, the proportion of dots drawing directions from the signal distribution was decreased and the proportion of dots drawing directions from the noise distribution was increased. The observer responded when the cinematogram changed in appearance from upward flow to local random motion. After this response, the proportion of signal continued to decrease for a random interval (up to 6 seconds). The process was then reversed, with the proportion of signal increasing until the subject responded that the upward flow had reappeared. This terminated the trial. For the second trial type, the stimulus sequence was reversed. We started with an initial percept of local random motion, and measured the points of transition to a perception of upward flow and back again to local random motion. These two different trial structures—signal-first and noise-first—were designed to produce different histories of directional exposure and thus reveal any perceptual effects dependent upon the history of stimulation. For both trial types, the signal to noise ratio changed slowly, with the proportion of dots sampling from each distribution changing by just two dots per frame. At the frame rate of our display, 10 Hz, it took a minimum of 25 seconds for the display to shift from complete dependence on one distribution to complete dependence on the other.

Observers used one eye to watch the centre of a circular display 16° in diameter. The dots were presented at twice threshold luminance against a dim background. Over the course of 5 sessions we made 100 measurements of the signal proportion at the two perceptual transitions for each of our three naive observers.

We repeated this experiment with two narrower ranges of signal distribution: 90° and 1°. Each of the three signal distributions generated a different history of exposure. We expected that, as the distribution narrowed, the signal would become

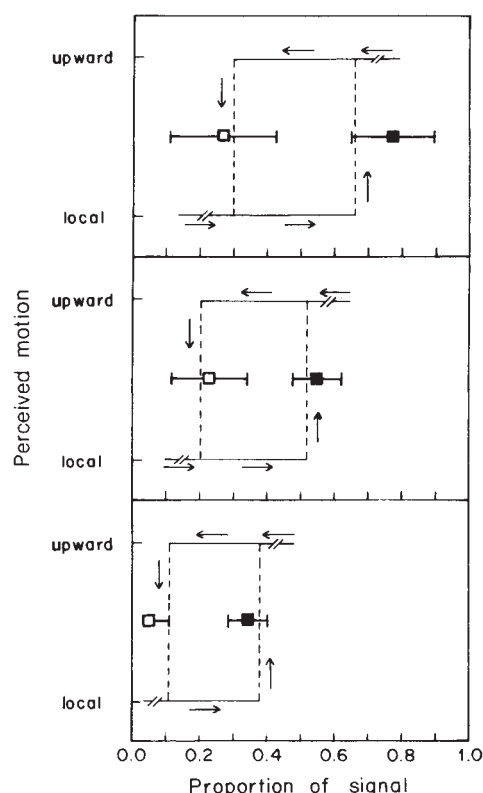


Fig. 1 Perceptual transitions measured under two different histories of stimulus exposure (indicated by arrows) for three signal distributions. The results, for observer J.F., were obtained for signal distributions whose ranges were: top, 180°; middle, 90°; bottom, 1°. Data points show proportion of 'signal' dots required for perceptual transition from local random motion to upward flow (■) and for perceptual transition from upward flow to local random motion (□). Error bars indicate one standard deviation (100 measurements). In each panel the separation between transition points measured with the different exposure histories is an index of hysteresis. The dashed lines mark the transition points calculated from a model incorporating cooperative interactions among direction-selective motion elements.

more effective in stimulating directionally selective visual elements that are tuned to upward motion. As a consequence, the occurrence of a transition between perceived states would require fewer signal dots.

The results for one observer are shown in Fig. 1; similar results were obtained for the other observers. Each panel shows data for a different range of signal. Note that the transitions differ for the three signal ranges. More importantly, the transition from noise-to-upward flow requires a larger number of dots sampled from the signal distribution than does the transition from upward flow-to-noise. For all three signal distributions, the two types of transition occur at significantly different ratios of signal to noise ($P < 0.005$).

Before attributing our results solely to neural hysteresis, a number of alternative explanations were considered and rejected. For example the motion after-effect, or waterfall illusion, did not produce our results as this after-effect would probably have facilitated, rather than retarded, the transition from upward motion to noise. Also, the results were largely unchanged by the use of a fixation point, suggesting that eye movements probably played little or no role. Finally, reaction time could be dismissed given the slow time course for changing the signal proportion.

Before developing a simple network that could describe our results, we needed a fuller account of how spatial parameters

might affect hysteresis. Using a 180° signal distribution, we made measurements (1) with a four-fold decrease in the spatial density of the cinematogram's dots, (2) with a four-fold decrease in the area of the display and (3) with the display shifted horizontally into the periphery of the visual field so that the nearest dot was 4° from fixation. None of these variations produced results that differed greatly from our original measurements. Of course, extreme changes in these display parameters would undoubtedly alter the hysteresis characteristics. However, because of these data and demonstrations suggesting that spatial variables have little effect on the dot interactions responsible for motion in cinematograms^{8,9}, we chose not to treat space explicitly in developing a model network to describe our results.

The model comprises a set of direction-selective mechanisms covering all 360° of motion direction, with each direction-selective mechanism having a gaussian profile for directional sensitivity. Based on previous results from our laboratory, the half-amplitude half-bandwidth of each mechanism's gaussian sensitivity profile was set to 30° (ref. 10). The model, whose mathematical formulation is a modification of the cooperative neural network previously proposed⁴, assumes non-linear excitatory interactions among mechanisms sensitive to similar directions of motion and non-linear inhibition among mechanisms sensitive to different directions. The dynamic response of this cooperative system is represented by a pair of coupled differential equations. For the excitatory activity, E_i , in direction channel i :

$$dE_i/dt = -E_i + (1 - E_i)S\left(P_i + \sum_j a_j E_j - \sum_j b_j I_j\right)$$

where S is a nonlinear function of sigmoidal shape, P_i is the external input to channel i , and a_j , b_j are the excitatory and inhibitory weights, respectively, of channel j with respect to channel i . A similar equation gives the inhibitory activity I_i , in channel i . In general terms, such interactions promote the formation of stable 'coalitions' among similarly tuned elements within the network¹¹. These neural coalitions can in turn produce various cooperative properties, including hysteresis.

The parameters of the model were constrained so that the model behaved in what is defined⁴ as the active transient mode. In this mode the system shows hysteresis, switching between different states of activity. In the model simulation the perception of local random motion is represented by a steady state of uniform activity across all mechanisms. Global upward flow is represented by a steady state in which the activity is localized about the mechanism selective for upward movement. A transition point is defined by the proportion of signal at which the network switches between these two states of activity. The results from the model are shown in Fig. 1. Dashed lines mark the transition points calculated from the model using a single parameter set. It can be seen that the model captures both the leftward shift and narrowing of the hysteresis profile with decreasing signal range. The model's behaviour here is readily explained. With decreasing signal range more activity is concentrated in fewer motion-selective elements arrayed about the upward direction. Consequently, a smaller proportion of signal dots suffices to indicate upward movement. In addition, fewer active elements reduces the opportunity for cooperative interactions in the network, which translates into a narrowing of the hysteresis profile².

If we accept the idea that the perception of motion direction depends upon a cooperative neural network, there may be circumstances outside the laboratory in which this network's cooperativity might be especially useful. Observers must extract a mean direction vector from a scene containing a great many different local vectors in many naturally-occurring situations. For example, we can judge the average direction in which the ocean's surf moves despite the fact that individual waves move along somewhat different paths; or we can judge the average

direction in which the wind blows the leaves of a tree despite the random variation of that movement from one leaf to the next (or for any one leaf over time). Faced with such multivectional stimuli, a cooperative network like the one embodied in our model would enhance the signal-to-noise ratio and thereby facilitate the perception of the mean direction of motion.

Supported by a grant from the US Air Force Office of Scientific Research.

Received 12 May; accepted 2 September 1986.

1. Julesz, B. *Am. Sci.* **62**, 32-43 (1974).
2. Fender, D. H. & Julesz, B. *J. Opt. Soc. Am.* **57**, 819-830 (1967).
3. Sperling, G. *Am. J. Psychol.* **83**, 461-534 (1970).
4. Wilson, H. R. & Cowan, J. *Kybernetik* **13**, 55-80 (1973).
5. Marr, D. & Poggio, T. *Science* **194**, 283-287 (1976).
6. Mayhew, J. E. W., Frisby, J. P. & Gale, P. *Perception* **6**, 207-208 (1977).
7. Chang, J. J. & Julesz, B. *Vis. Res.* **24**, 1781-1788 (1984).
8. Williams, D. W. & Sekuler, R. *Vis. Res.* **24**, 55-62 (1984).
9. Baker, C. L., Jr. & Braddick, O. J. *Vis. Res.* **22**, 1253-1259 (1982).
10. Williams, D. W., Twetten, S. & Sekuler, R. *Invest. Ophthalmol. vis. Sci. Suppl.* **25**, 14 (1984).
11. Feldman, J. A. & Ballard, D. H. in *Human and Machine Vision* (eds Beck, J., Hope, B. & Rosenfeld, A.) 107-155 (Academic, New York, 1983).

Inhibitory neurones of a motor pattern generator in *Xenopus* revealed by antibodies to glycine

N. Dale*†, O. P. Ottersen‡, Alan Roberts§ & J. Storm-Mathisen‡||

‡ Anatomical Institute, University of Oslo, Karl Johans gate 47, N-0162 Oslo 1, Norway

* Department of Physiology III, Karolinska Institute, Lidingövägen 1, Stockholm S-114 33, Sweden

§ Department of Zoology, University of Bristol, Woodland Road, Bristol BS8 1UG, UK

Glycine and γ -aminobutyric acid (GABA) are inhibitory transmitters of major importance^{1,2}. Whereas neurones using GABA as the transmitter can be visualized by immunocytochemical methods for glutamate decarboxylase (GAD)^{3,4} or GABA⁵⁻⁸, no comparable techniques have been available for the selective visualization of glycinergic neurones. We have now produced polyclonal antibodies which specifically recognize glycine in glutaraldehyde-fixed tissue. We used these antibodies to investigate the distribution of glycine in the simple central nervous system (CNS) of the *Xenopus* embryo, which contains an anatomically and physiologically defined class of reciprocal inhibitory interneurons, the commissural interneurons⁹⁻¹¹. These interneurons have an important role in the generation of the swimming motor pattern and are thought to be glycinergic¹¹⁻¹³. The glycine antibodies specifically stain these interneurons, revealing their distribution and number in the embryo CNS. This is the first demonstration of the selective localization of glycine-like immunoreactivity in a putative glycinergic class of neurone that has been characterized physiologically, pharmacologically and anatomically.

Evidence for glycine as a transmitter in the vertebrate spinal cord has been based on four types of information: the actions of exogenous glycine^{14,15}, specific blockade by the antagonist strychnine^{14,15}, the depletion of glycine following ischaemic degeneration of spinal interneurons^{16,17} and the demonstration of neuronal uptake of ³H-glycine¹⁸⁻²¹. However, antagonism of an inhibitory postsynaptic potential (i.p.s.p.) by strychnine is not unequivocal evidence that glycine is the endogenous transmitter: the hyperpolarizing actions of β -alanine and taurine (amino acids closely related to glycine; Fig. 1) are also blocked by strychnine². In the absence of further evidence β -alanine

and taurine must therefore also be considered as candidate transmitters. We decided to raise antibodies which could selectively recognize glycine in glutaraldehyde-fixed nervous tissue and could be used to localize glycine in neurones. Modifications of our methods^{5,6} have been used previously by Pourcho and Goebel²² to produce and test a glycine antiserum which marks cells in the cat retina and brain stem.

Glycine was coupled to carrier proteins by means of glutaraldehyde as described previously⁵. As suggested by Seguela *et al.*⁷, the carrier protein was changed, here for each immunization, to avoid boosting the immune response to the carrier protein and any particular configurations of the glycine-glutaraldehyde complex that might depend on the protein (Fig. 1 legend). Unwanted antibodies recognizing glutaraldehyde-treated amino acids and proteins were removed by immunosorbent chromatography on a sequence of columns (Fig. 1 legend). The purified antiserum was initially tested for crossreactivity with a series of compounds conjugated by glutaraldehyde to a total macromolecular extract from rat brain and spotted on cellulose ester filters⁶. The compounds were GABA, taurine, phosphoethanolamine, β -alanine, L- α -alanine, valine, leucine, proline, serine, threonine, tryptophan, tyrosine, histidine, lysine, arginine, ornithine, aspartate, N-acetylaspartate, glutamate, asparagine, glutamine, methionine, cysteine, reduced glutathione, α -aspartylglycine, carnosine, homocarnosine, cadaverine, putrescine, spermine and spermidine. While the glycine conjugate was intensely stained, none of the other conjugates were stained appreciably darker than 'null' conjugates (macromolecular extract treated with glutaraldehyde in the absence of amino acid). The staining was abolished by solid-phase absorption with glycine fixed to ovalbumin or to lysine by glutaraldehyde (Fig. 1 legend).

To check the specificity of these antibodies in the *Xenopus* embryo we made two different sets of conjugates, one of macromolecules from whole embryos and one of macromolecules from the CNS of adult *Xenopus*. To ensure identical conditions for testing and immunocytochemistry, the test filters with these conjugates were processed in the same vessels as the tissue preparations. The tissue consisted of the CNS of *Xenopus* embryos stage 29/30 to 37/38 (ref. 23) raised by induced breeding in Bristol, fixed in glutaraldehyde and shipped to Oslo for dissection and processing⁶ of the CNS.

The results with *Xenopus* conjugates (Fig. 1a) were similar to those with rat brain conjugates. A glycine-glutaraldehyde complex, added to the diluted antiserum (at a final concentration of 100 or 300 μ M with respect to glycine) completely abolished staining of the glycine spots and of the tissue (Fig. 1e,f). Complexes of β -alanine (Fig. 1c,d), L- α -alanine, GABA or reduced glutathione (all at 300 μ M), or a mixture of individual glutaraldehyde complexes of taurine, aspartate, glutamate, glutamine and serine (all at 300 μ M) produced no reduction in the staining of glycine test spots (the converse is seen in ref. 32) or tissue.

The *Xenopus* embryo spinal cord appears on the basis of retrograde horseradish peroxidase (HRP) filling to contain eight anatomically distinct neurone classes⁹. One of these, the commissural interneurons, is defined by the following anatomical features⁹: (1) a unipolar soma in a middle to dorsal position in the cord, (2) a single ventral process bearing dendrites that extend laterally into the marginal zone (lateral tract) and (3) an axon crossing ventrally to ascend and sometimes also descend contralaterally (Fig. 3a). These neurones have been shown by intracellular recording and HRP injection to be active during fictive swimming¹⁰ and to produce mid-cycle i.p.s.ps in contralateral motoneurons¹¹. The mid-cycle i.p.s.ps in motoneurons and the response to exogenous glycine (100 μ M) have a reversal level near the resting potential, depend on an increase in chloride permeability and are unaffected by 10 μ M bicuculline, but are blocked by 1 μ M strychnine^{11,13,24}. The commissural interneurons may therefore be glycinergic.

† Present address: Howard Hughes Medical Institute, Center for Neurobiology and Behavior, Columbia University, 722 West 168th Street, New York 10032, USA.

|| To whom correspondence should be addressed.

Fig. 1 The purified glycine antiserum (31, 1:90) gives selective staining of test spots of glycine fixed by glutaraldehyde to macromolecules from *Xenopus* embryos (pigmented) (○) or adult CNS (○) (a, i), and selective staining of a characteristic class of neurones, the commissural interneurons, in whole-mounts of the spinal cord of *Xenopus* embryos (lateral view) (b). The specific staining is abolished by glutaraldehyde complexes of glycine (Gly-G) (e, f), but not by complexes of the closely related β -alanine (300 μ M added to the diluted serum) (c, d). In comparison, a purified GABA antiserum (26, 1:200)⁶ stains only the GABA spots (g) and different cells²⁵ (h) from those stained by serum 31. The structures of the amino acids most closely related to glycine are shown (j); standard symbols are used for amino acids; GSH, γ -glutamylcysteinylglycine (glutathione).

Methods. The serum was raised by immunizing a rabbit (intracutaneously on the back at 6-week intervals with 125 μ g protein in Freund's complete adjuvant; Gibco) with glutaraldehyde conjugates (0.3–1 μ mol of glycine bound per mg protein) of a sequence of carrier proteins: keyhole limpet haemocyanin (Calbiochem), casein (Sigma), human haemoglobin (Sigma), lysozyme (Worthington), *Limulus polyphemus* haemocyanin (Sigma), horse skeletal muscle myoglobin (Sigma) and bovine thyroglobulin (Sigma). Portions (50 μ l) of the crude serum were applied to a set of three microcolumns (i.d. 0.3 cm) connected in series and containing protein (6 mg per ml gel) coupled to agarose particles (Sephacrose 4B, Pharmacia²⁷) and reacted with glutaraldehyde (Taab, EM-grade, 75 μ mol per ml gel) in the presence or absence of amino acid (20 μ mol per ml gel). Column 1 carried bovine serum albumin treated with glutaraldehyde, column 2 carried albumin treated with glutaraldehyde in the presence of glutamate, and column 3 consisted of portions of agarose particles coupled with the individual carrier proteins and treated with glutaraldehyde. The serial column was eluted with Tris buffer (0.1 M pH 7.4) containing NaCl (0.3 M). Glycine buffer (0.1 M pH 2.3 with 0.3 M NaCl) or propionic acid (2 M) was then used to regenerate the column. Columns 1 and 2 were essential for specificity, column 3 did not change the specificity of staining of test conjugates or tissue, but slightly reduced the background. Absorption of the purified serum with mycobacteria from the adjuvant did not change the staining of test conjugates or tissue. The specific antibodies could be completely removed by passage through a microcolumn of agarose particles carrying glycine fixed to ovalbumin by glutaraldehyde, or preformed complexes of glycine and glutaraldehyde (see below) attached to lysine-Sepharose (Pharmacia). CNS of *Xenopus* embryos (fixed in 5% glutaraldehyde diluted from 25% in 0.1 M sodium phosphate buffer pH 7.4 (NaP_i)) and test filters⁶ were incubated together in the purified serum (dilution 1:90–1:400, 3–6 days, 4–6 °C) with 0.5% Triton X-100 and further processed essentially as described previously⁶. Penetration of reactants was enhanced by soaking the CNS in sucrose (30% in NaP_i), freezing in liquid N₂ and thawing at room temperature before incubation with the antisera. The test conjugates spotted on the filter semicircles were prepared by mixing 4.5 mg tissue protein per ml in NaP_i with 20 mM amino acid and 50 mM glutaraldehyde for 1 h, and dialysing. Specificity was further demonstrated by adding glutaraldehyde complexes of various amino acids to the diluted serum. The complexes were prepared by mixing amino acid (20 μ mol) with glutaraldehyde (40 μ mol, that is, optimal stoichiometry²⁸, total volume 116 μ l) for 1 h at room temperature and adding ethanolamine HCl (1 M 384 μ l), all in NaP_i (see also ref. 31).

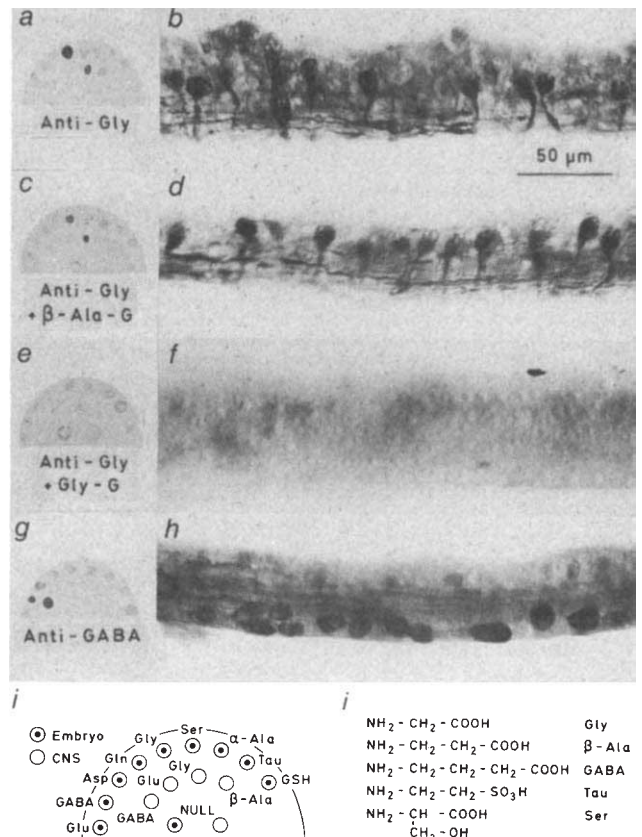
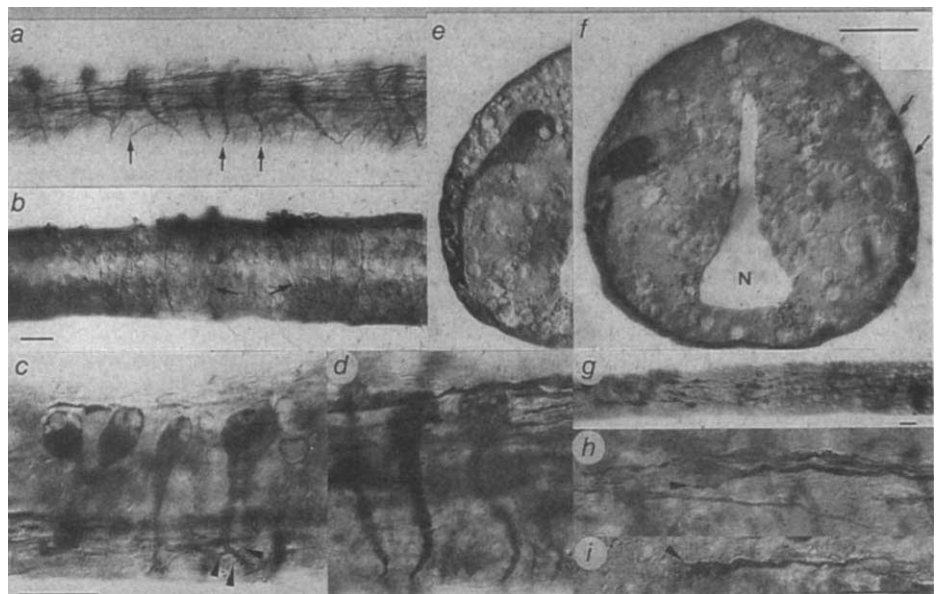


Fig. 2 Characteristic anatomical features of the commissural interneurons of *Xenopus* embryo spinal cord (see also Fig. 3) revealed by specific staining with the glycine antiserum (see Fig. 1 legend). a, Lateral view of a slightly rotated spinal cord showing unipolar somata with axons (arrows) crossing ventrally under the neurocoel. b, Ventral view of the spinal cord showing the crossing axons (arrows). c, d, High-power lateral view of stained cells with dendrites (arrowheads) coming out of the fat proximal process. e, f, Transverse sections (6 μ m) of immunostained spinal cords osmicated and embedded in Araldite showing two unipolar somata with axons crossing ventrally under the neurocoel (N). The marginal zones clearly contain transversely cut stained axons (arrows). g-i, Stained axons and growth cones in the lateral axon tracts. Arrowheads show short micropodia emanating from growth cones. The large unstained inclusions are yolk platelets which are present in all the embryonic cells. Interference contrast optics; all from whole-mounts except e, f. Scale bars, 20 μ m.



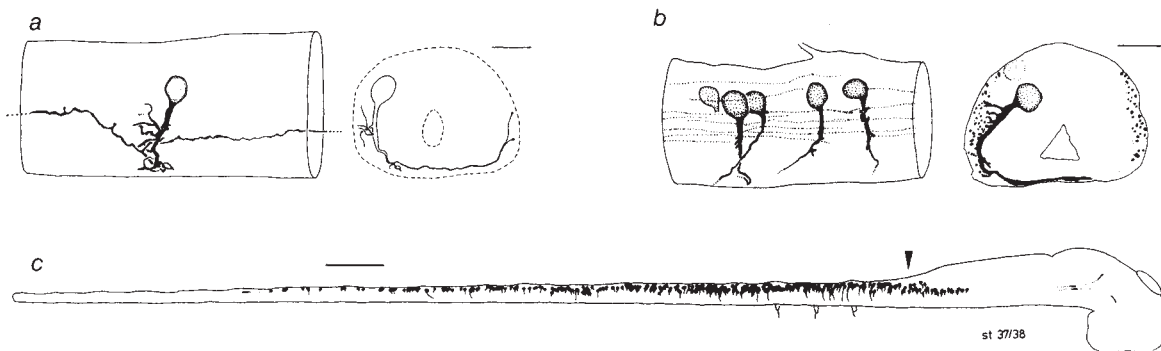


Fig. 3 Camera lucida drawings of commissural interneurons from *Xenopus* embryo spinal cord. *a*, Commissural interneurone filled with horseradish peroxidase from the intracellular recording micropipette in a stage 37/38 embryo. Left, lateral view; right, reconstruction of the transverse section view of these cells (based on micrometer readings in the side view) showing dendrites and axon passing under cord before forming ascending and descending branches (redrawn from ref. 10). *b*, Neurones stained with the glycine antiserum showing similarity with *a*. Left, drawing of trunk spinal cord from whole-mount at stage 33/34 (lateral view). Four of the stained cells have a unipolar soma with their main ventral process with short dendrites, and an axon which runs ventrally to cross to the other side. Dotted lines show stained axons. Right, 6- μ m transverse plastic section of stage 37/38 rostral trunk cord. Stained axons are present throughout the marginal zone on each side. One faint and one clear soma are present on the left side. The latter has a large ventral process giving off dendrites into marginal zone and continuing ventrally to run under neurocoel to the opposite side. *c*, Selective distribution of glycine-positive neurones along the neuraxis in a stage 37/38 embryo. Arrowhead shows border between hindbrain and spinal cord. All neurones (299) on one side are shown. Scale bars, *a*, *b*, 25 μ m; *c*, 200 μ m.

To investigate whether these neurones contain an elevated glycine concentration, we processed the embryo CNS with the glycine antibodies. This gave specific staining of a population of cells with the three defining features of commissural interneurons outlined above (Figs 2*a-f*, 3*b*). At stage 37/38²³ there were about 300 cells per side extending from the mid-hindbrain to the caudal spinal cord (Fig. 3*c*). Comparison with results from retrograde HRP filling indicated that the glycine antiserum had stained the whole population of commissural interneurons. The most caudal somata were often paler than the rest, suggesting that cells in this region were still differentiating. A few of the stained somata lay close to the cord surface in a dorsolateral position (Fig. 3*b*). These cells may either be commissural interneurons with slightly different anatomy, or members of the other class of spinal interneurone with crossed axons, the dorso-lateral commissural interneurons⁹. Staining of ventral commissural and longitudinal axons was darker than that in somata. Some axons bifurcated before reaching the other side. Ascending axons reached the rostral hindbrain while a few extended into the midbrain or even the forebrain. Descending axons reached to within 600 μ m of the caudal end of the spinal cord at stage 37/38. Many axons terminated in large, stained growth cones (Fig. 2*g-i*). No other groups of somata were stained.

An antiserum specific for GABA⁶ stained spinal interneurons with ascending axons and one class of possible receptor cells, the ciliated ependymal cell (Fig. 1*h*)^{9,25}. However, it did not stain any commissural interneurons or ventral commissural axons in the spinal cord. Antisera specific for taurine²⁶ or β -alanine (S. Davanger, A. Bruun, B. Ehinger, J. Laake, L. Orensanz, O.P.O. & J.S.-M., in preparation) produced no staining of cells or fibres in the *Xenopus* embryo spinal cord. Our immunocytochemical data together with the pharmacological and physiological evidence thus suggest that glycine rather than another inhibitory amino acid is the transmitter of the commissural interneurons.

We thank A. T. Bore, T. Eliassen, I. Fridström, G. F. Lothe, K. Stentoft, L. J. Teagle and J. Line Vaaland for assistance. This work was supported by the Norwegian Research Council for Science and the Humanities, the SERC, the Company of Biologists, the Norwegian Council on Cardiovascular Disease and the NfTKB.

Note added in proof: Campistrone *et al.*²⁹ have recently reported on an antiserum to glutaraldehyde-fixed glycine (reacted with NaBH₄) and used it to stain neurones in rat brain and spinal cord. With a few exceptions, their results are similar to those obtained in rat with our antiserum³⁰.

Received 30 May; accepted 28 July 1986.

- Curtis, D. R. & Johnston, G. A. R. *Ergebn. Physiol.* **69**, 97-188 (1974).
- Krnjević, K. *Physiol. Rev.* **54**, 418-540 (1974).
- Saito, K. *et al. Proc. natn. Acad. Sci. U.S.A.* **71**, 269-273 (1974).
- Mugnaini, E. & Oertel, W. *Handbook of Chemical Neuroanatomy* Vol. 4 (eds Björklund, A. & Hökfelt, T.) (Elsevier, Amsterdam, 1985).
- Storm-Mathisen, J. *et al. Nature* **301**, 517-520 (1983).
- Ottersen, O. P. & Storm-Mathisen, J. *J. comp. Neurol.* **229**, 374-392 (1984).
- Seguela, P., Geffard, M., Buijs, R. M. & Le Moal, M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 3888-3892 (1984).
- Hodgson, A., Penke, B., Erdei, A., Chubb, I. W. & Somogyi, P. *J. Histochem. Cytochem.* **33**, 229-239 (1985).
- Roberts, A. & Clarke, J. D. W. *Phil. Trans. R. Soc. B* **296**, 195-212 (1982).
- Soffe, S. R., Clarke, J. D. W. & Roberts, A. *J. Neurophysiol.* **51**, 1256-1267 (1984).
- Dale, N. *J. Physiol., Lond.* **363**, 61-70 (1985).
- Roberts, A., Soffe, S. R. & Dale, N. in *Neurobiology of Vertebrate Locomotion* (eds Grillner, S. *et al.*) (Macmillan, London, 1986).
- Soffe, S. R. *J. Physiol., Lond.* **382** (1987).
- Werman, R., Davidoff, R. A. & Aprison, M. H. *Nature* **214**, 681-683 (1967).
- Curtis, D. R., Hösli, L. & Johnston, G. A. R. *Expl Brain Res.* **6**, 1-18 (1968).
- Davidoff, R. A., Graham, L. T. Jr, Shank, R. P., Werman, R. & Aprison, M. H. *J. Neurochem.* **14**, 1025-1031 (1967).
- Homma, S., Suzuki, T., Murayama, S. & Otsuka, M. *J. Neurochem.* **32**, 691-698 (1979).
- Hökfelt, T. & Ljungdahl, Å. *Brain Res.* **32**, 189-194 (1971).
- Ljungdahl, Å. & Hökfelt, T. *Brain Res.* **62**, 587-595 (1973).
- Matus, A. I. & Dennison, M. E. *Brain Res.* **32**, 195-197 (1971).
- Price, D. L., Stocks, A., Griffin, J. W., Young, A. & Peck, K. *J. Cell Biol.* **68**, 389-395 (1976).
- Pourcho, R. G. & Goebel, D. J. *Brain Res.* **348**, 339-342 (1985).
- Nieuwkoop, P. D. & Faber, J. *Normal Tables of Xenopus laevis* (Daudin) (North-Holland, Amsterdam, 1956).
- Soffe, S. R. & Roberts, A. *J. Neurophysiol.* **48**, 1279-1288 (1982).
- Dale, N., Roberts, A., Ottersen, O. P. & Storm-Mathisen, J. (in preparation).
- Madsen, S., Ottersen, O. P. & Storm-Mathisen, J. *Neurosci. Lett.* **60**, 255-260 (1985).
- Porath, J., Aspberg, K., Drevin, H. & Axén, R. *J. Chromatogr.* **86**, 53-56 (1973).
- Hardy, P. M., Nicholls, A. C. & Rydon, H. N. *J.C.S. Perkin Trans.* **9**, 958-962 (1976).
- Campistrone, G., Buijs, R. M. & Geffard, M. *Brain Res.* **376**, 400-405 (1986).
- Ottersen, O. P., Davanger, S. & Storm-Mathisen, J. *expl Brain Res.* (in the press).
- Ottersen, O. P., Storm-Mathisen, J., Madsen, S., Skumlien, S. & Strømhaug, J. *med. Biol.* **64**, 147-158 (1986).

T-cell epitope of the autoantigen myelin basic protein that induces encephalomyelitis

Scott S. Zamvil*, Dennis J. Mitchell*, Anne C. Moore*, Kumiko Kitamura*, Lawrence Steinman* & Jonathan B. Rothbard†‡

* Departments of Neurology and Pediatrics, Stanford University, Stanford, California 94305, USA

† The Imperial Cancer Research Fund, P.O. Box 123, Lincoln's Inn Fields, London WC2A 3PX, UK

Chronic relapsing paralysis and demyelination within the central nervous system (CNS), features associated with the human disease multiple sclerosis (MS)¹⁻⁴, develop in mice after injection of murine T-cell clones specific for the autoantigen myelin basic protein (MBP)^{5,6}. We examined the fine specificity of three independently derived encephalitogenic T-cell clones using synthetic polypeptides derived from portions of the N-terminal sequence of MBP. These clones appear functionally identical; they all respond to an epitope in the N-terminal nine amino acid residues in association with the same class II (I-A) molecules of the major histocompatibility complex (MHC). Both the N-terminal acetyl moiety and the first residue (Ala) are necessary for recognition. Only N-terminal MBP peptides recognized by these clones were found to cause encephalomyelitis (EAE) *in vivo*. These results show that the N-terminal MBP-specific T lymphocytes that mediate autoimmune encephalomyelitis are a small population with a limited repertoire; they all recognise the same combination of MHC and target.

EAE is a model for autoimmune diseases mediated by class II restricted T lymphocytes⁵⁻¹⁰. Immunization with intact MBP or fragments of MBP causes EAE in susceptible mouse strains¹⁰⁻¹². However, not all such fragments are encephalitogenic. Only the N-terminal 37 amino-acid MBP fragment causes EAE in PL/J (H-2^b) and (PL/J × SJL/J (H-2^d))F₁ ((PLSJ)F₁) mice^{11,12}. Encephalitogenic T-cell clones with the L3T4⁺, Lyt2⁻ phenotype have been isolated from both these strains following immunization with either intact rat or bovine MBP^{5,6}. Clones PJR-25 and F1-12 were isolated from homozygous PL/J and (PLSJ)F₁ mice, respectively, after immunization with intact rat MBP. Encephalitogenic clone PJB-20 was isolated from homozygous PL/J mice after immunization with intact bovine MBP. These clones respond to a determinant shared with mouse (self) MBP within the encephalitogenic N-terminal MBP fragment in association with I-A^b(A_α^bA_β^b) class II molecules. Class II-restricted T lymphocytes recognize discrete, contiguous sequences within proteins, sometimes fewer than ten residues¹³ in length via their antigen-specific cell surface receptors^{14,15}. To investigate the pathogenesis of this autoimmune disease further, we examined whether these encephalitogenic T lymphocytes all respond to a common epitope, or have distinct N-terminal MBP specificities.

Separate forms of intact MBP with different N-terminal sequences were tested for their ability to stimulate proliferation of encephalitogenic T-cell clones. Bovine MBP was found to be less stimulatory than mouse or rat MBP at the same concentration (results not shown). Since the N-terminal 37 amino-acids of bovine MBP differ from mouse MBP only at residues 2 and 17 (see Fig. 1)¹⁶, we predicted that the epitope(s) recognized by these T-cell clones would include one of these two residues. The overlapping peptides pR(rat)1-16 and pR5-20 were made and tested for their ability to stimulate proliferation of these clones. The epitope recognized by all three clones is found in the N-terminal 16 residues; peptide pR1-16 is as effective as intact

		1								10									20
Rat/human MBP	Ac-	A	S	Q	K	R	P	S	Q	R	H	G	S	K	L	A	T	A	S
Mouse MBP	Ac-	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Bovine MBP	Ac-	A	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	S	—
pR 1-16	Ac-	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
pR 5-20									Ac-	—	—	—	—	—	—	—	—	—	—
pR 1-9	Ac-	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
pR 1-11	Ac-	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
pR 2-11									Ac-	—	—	—	—	—	—	—	—	—	—
pR ⁺ 1-11																			
pBR 1-11	Ac-	A	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—

Fig. 1 Amino-acid sequences of MBP from different species, and synthetic MBP peptides. The sequences of intact rat, human, mouse, and bovine MBP have been published previously¹⁶. Peptides corresponding to the sequences of rat (R) and bovine (B) MBP were synthesized as described previously^{29,30} by solid phase techniques on a Beckman Model 990 B peptide synthesizer; using commercially available amino acid polystyrene resins and t-Boc-protected amino acids from Peninsula Laboratories. The protecting groups used were: o-benzyl esters for Thr and Ser; tosyl for His; orthochlorobenzylloxycarbonyl for Lys; and 2,6-dichlorobenzyl for Tyr. Couplings were performed using a 2.5 molar excess of t-Boc amino acid and dichlorohexylcarbodiimide (DCC) over the number of milliequivalents of amino acid on the resin. All couplings were >99% complete, as determined by reaction with ninhydrin. The peptides were deprotected and removed from the resin simultaneously by treatment with anhydrous hydrogen fluoride in the presence of anisole, dimethylsulphide, and indole, then separated from organic side products by extraction with ether and isolated from the resin by extraction with 5% acetic acid and subsequent lyophilization. Purity was determined by high pressure liquid phase column (Merck) and amino acid analysis. The peptides used in these experiments were not further purified as they all contained >90% of the desired peptide.

MBP (Fig. 2a) whereas pR5-16 and pR5-20 do not stimulate proliferative responses. These clones also respond to peptides pR1-9 and pR1-11. Since pR1-11 and pR1-16 are equipotent, residues 12-16 do not contribute significantly to the epitope recognized by these T-cell clones.

Acetylation of the first residue (N-Ac-Ala₁) is essential for recognition. When the acetyl group is removed, revealing a positively charged amino terminus (pR⁺1-11, Fig. 1), proliferative activity is eliminated (Fig. 2b). The amino-acid Ala₁ is also necessary, since the MBP peptide pR2-11, containing N-Ac-Ser₂ is not stimulatory even at the highest concentrations tested. Bovine MBP, which has an alanine instead of a serine residue at position 2 (see Fig. 1), is less encephalitogenic than intact rat or mouse MBP, which both have Ser₂, in homozygous PL/J and (PLSJ)F₁ mice¹¹. Peptide pBR1-11, which has Ala at the second residue in a rat MBP1-11, sequence (Fig. 1) is less effective than pR1-11 and requires higher concentrations to stimulate proliferation. Thus, these representative T-cell clones isolated after stimulation with rat MBP or bovine MBP share the same N-terminal specificity. We also identified a second T-cell epitope in the encephalitogenic N-terminal fragment of rat MBP. T-cell clones that are restricted to A_α^bA_β^b class II molecules isolated from (PLSJ)F₁ mice recognize rat but not mouse (self) MBP, and are not encephalitogenic. The epitope recognized by these T-cell clones is found in pR9-16 (Fig. 3). Thus there are two distinct epitopes in the N-terminal 16 residues of MBP.

MBP peptide fragments have been shown to cause acute encephalomyelitis in other animal models¹⁷⁻²². However, peptides as small as those described in this report have not previously been shown to cause EAE in mice. Peptides pR1-9, pR1-11, and pR1-16, all of which are recognized *in vitro* by encephalitogenic T cell clones, cause EAE in (PLSJ)F₁ mice when given in adjuvant, without carrier protein (Table 1). Histo-

‡ To whom correspondence should be addressed.

Table 1 Induction of EAE by synthetic N-terminal MBP peptides

Antigen	Dose (nanomoles)	Incidence	Mean severity	Mean day of onset
pR1-9	100 (110 µg)	10/10	4.2 (±0.63)	11.9 (±0.74)
pR1-9	20	4/5	1.8 (±0.96)	12.5 (±1.7)
pR1-11	100 (128 µg)	10/10	4.4 (±1.3)	12.8 (±1.4)
pR1-11	20	3/5	3.7 (±1.2)	12.3 (±1.2)
pR1-16	100 (186 µg)	3/5	1.7 (±0.57)	14.3 (±1.5)
pR1-16	20	3/5	2.7 (±1.2)	13.7 (±1.2)
pR2-11	200 (244 µg)	0/10	—	—
pR2-11	100	0/5	—	—
pR2-11	20	0/5	—	—
pR5-16	100 (144 µg)	0/5	—	—
pR9-16	100 (97 µg)	0/5	—	—
pR11-27	100 (184 µg)	0/5	—	—
pR20-37	100 (216 µg)	0/5	—	—

MBP peptides were tested for their ability to cause EAE following the protocol described previously^{11,12} for induction of EAE in mice with intact MBP. Each peptide was dissolved in phosphate buffered saline (PBS) and emulsified with complete Freund's adjuvant (CFA) in a 1:1 mixture of PBS:CFA containing 4 mg ml⁻¹ H37Ra (Difco). Recipient (PLSJ)F₁ female mice (Jackson Laboratories, Maine), aged 12–20 weeks, were injected with 0.1 ml emulsion at the base of the tail. On the same day 10¹⁰ heat killed *Bacillus pertussis* (Michigan Department of Health, Lot 91B) were injected intravenously. Forty-eight hours later 10¹⁰ *B. pertussis* were injected a second time. Mice were examined daily for signs of EAE by two independent observers. Severity of EAE was graded using a scale previously described⁶: 0, no sign of EAE; 1, decreased tail tone only; 2, mild paraparesis; 3, moderately severe paraparesis; 4, complete paraplegia; 5, moribund. Synthetic peptides pB1-9, pBR1-11, and pR5-20 have also been tested (results not shown). When (PLSJ)F₁ mice are injected with 100 µmoles pB1-9 or pBR1-11, peptides that are recognized by encephalitogenic T cell clones, clinical and histological EAE are observed. Peptide pR5-20 is not encephalitogenic.

logical examination of paralysed mice shows perivascular infiltration of mononuclear cells in the CNS, as observed in EAE caused by immunization with whole MBP^{10,11}. In contrast, peptides pR2-11, pR5-16, pR9-16, pR11-27 and pR20-37, which do not stimulate encephalitogenic T-cell clones to proliferate *in vitro*, do not cause clinical or histological signs of EAE in (PLSJ)F₁ mice (Table 1).

T-cell clones recognize peptide sequences in association with a particular class II MHC product. Encephalitogenic T-cell clones isolated after immunization with intact MBP recognize pR1-11 in association with I-A^u class II molecules (Fig. 2b). It is possible that encephalitogenic (PLSJ)F₁ T lymphocytes may recognize this epitope in association with other class II molecules. However, immunization with pR1-11 causes EAE in H-2^u strains (PLJ 7/10 mice paralysed and B10.PL 10/10 paralysed), but not H-2^s strains (SJL/J 0/10 paralysed and B10.S 0/10 paralysed), and lymphocytes primed *in vivo* with pR1-11 are inhibited from responding *in vitro* to either pR1-11 or intact MBP by monoclonal antibodies specific for I-A, but not I-E molecules (results not shown). Thus, the encephalitogenic responses to intact MBP and N-terminal MBP peptides in PL/J and (PLSJ)F₁ mice are restricted to I-A^u class II molecules.

We have thus identified the epitope recognized by T-cell clones which mediates autoimmune encephalomyelitis. These independently-derived T-cell clones appear to have functionally identical peptide specificity and class II restriction. Further, the induction of autoimmune encephalomyelitis by *in vivo* administration of N-terminal MBP peptides correlates with T cell recognition *in vitro*; of all N-terminal MBP peptides tested, only those peptides which are recognized by these independently-derived T-cell clones are encephalitogenic. T-cell clones derived after immunization with pBR1-11 or pR1-11 also have the same N-terminal MBP and class II specificity as the clones

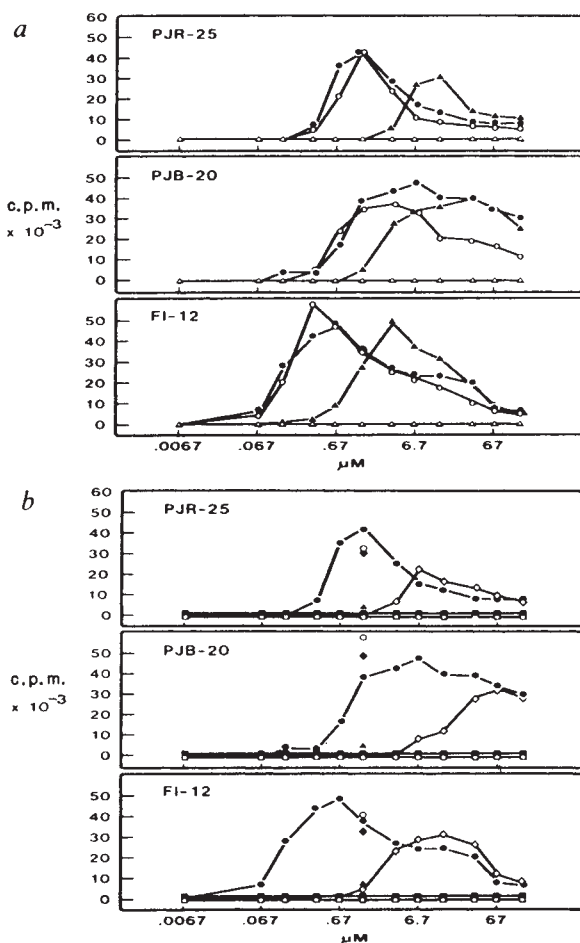


Fig. 2 Recognition of N-terminal MBP peptides by encephalitogenic T cell clones. **a**, Peptides pR1-16 (○), pR1-11 (●), pR1-9 (▲), and pR5-20 (△) were tested for their ability to stimulate encephalitogenic proliferative T cell clones, isolated as described previously^{1,6,31}. To assay for proliferative responses, 10⁴ T cells were cultured with 5 × 10⁵ X-irradiated (3,000 rad) PL/J splenic antigen presenting cells (APC) in 0.2 ml culture media in 96 well flat-bottomed microtitre plates (Falcon). Culture medium was RPMI 1640 (Gibco; supplemented with 10% fetal calf serum (Hyclone), 5 × 10⁻⁵ M 2-mercaptoethanol, 2 mM glutamine, 100 U ml⁻¹ penicillin and 100 µg ml⁻¹ streptomycin). Peptides were added to cultures giving the final concentrations indicated. At 48 h each well was pulsed with 1 µCi ³H-thymidine and cells collected 16 h later. The mean c.p.m. for thymidine incorporation was calculated for triplicate cultures. Standard deviations are within 10% of the mean. The same pattern of response to N-terminal peptides is found when these T cell clones are cultured with (PLSJ)F₁ APC. **b**, Peptides pR1-11 (○), pR1-11 (□), and pR2-11 (■) were tested as described above. In a separate assay, the proliferative response to pR1-11 (○) was inhibited with 1 µg per well anti I-A monoclonal antibody 10-2.16 (ref. 32) (▲), but not anti I-E monoclonal antibody 14-4-4³³ (◆). When tested at this concentration, 14-4-4 inhibits T cell clones restricted to I-E molecules⁶ (results not shown).

described here. Clones derived from mice immunized with pBR1-11 respond in a heteroclitic manner; like PJB-20 they respond better to pR1-11 than to pBR1-11, the original immunogen (results not shown). Thus, *in vitro* and *in vivo* results suggest that the encephalitogenic T-cell response to MBP 1-37 is limited to a discrete population of MBP-reactive T cells.

It has recently become possible to examine the genes that encode antigen-specific T-cell receptors (TCR)²³⁻²⁵ and investigate the relationship between specificity and TCR gene expression. Ovalbumin-specific T-cell clones that recognize the same epitope and share the same class II restriction have been shown to use the same V_α/J_α and V_β/J_β TCR genes²⁶; T-cell clones that share the same fine specificity for cytochrome C use the same V region genes, but can express different J elements²⁷. However examination of TCR genes from clones of different specificities suggests that there may not be a simple correlation

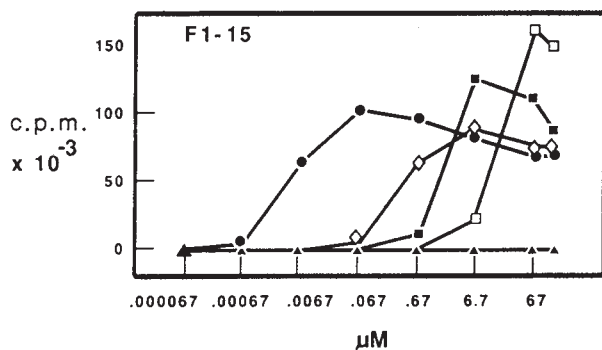


Fig. 3 N-terminal peptides recognized by a representative non-encephalitogenic MBP specific T cell clone. Peptides pR1-16 (●), pR5-16 (◇), pR7-16 (■), pR9-16 (□) and pR10-20 (▲), were tested as described (Fig. 2, legend) F1-15 T cells were cultured with (PLSJ)_{F1} splenic APC. These N-terminal MBP-specific non-encephalitogenic T cell clones are restricted to Aα^sAβ^u class II molecules. N-terminal peptides that are recognized by these non-encephalitogenic T cell clones do not cause EAE (see Table 1).

between V_β expression and T-cell specificity²⁸. The TCR genes from encephalitogenic T-cell clone PJR-25 are being cloned, and should be useful probes allowing us to examine whether other encephalitogenic clones, which share the same epitope and class II restriction, express the same TCR genes.

We have described here the first investigation of the fine specificity of T cell clones mediating an autoimmune disease. Two distinct T-cell epitopes in MBP 1-16 have been characterized. Examination of the response to MBP peptides, both *in vitro* and *in vivo*, has revealed that only a limited repertoire of N-terminal specific T lymphocytes, expressing a common class II restriction and peptide specificity, mediate induction of EAE. These results may provide insights into the immunogenetic mechanisms of other autoimmune diseases. In cases where only a limited repertoire of T lymphocytes mediate autoimmune pathogenesis, specific forms of therapy might be possible. For example measures which selectively inactivate or eliminate only autoaggressive T cells bearing the N-terminal MBP specificity may be beneficial in the treatment of autoimmune encephalomyelitis.

We thank Teri Montgomery for her valuable assistance and Professors Hans Acha-Orbea, Mark Davis, Robert Fritz, Garry Fathman, Phillipa Marrack, Hugh McDavitt, Patricia Nelson, Tim Vollmer and Matthew Waldor for helpful discussions. I.S. is supported by NIH grants NS00571, AI22462 and NS18235, the Rosenthal Foundation, the Fausel Foundation and the Swim Foundation.

Received 16 May; accepted 6 August 1986.

1. Wisniewski, H. M. & Keith, A. B. A. *Neurol.* **1**, 144-148 (1977).
2. Raine, C. S., Snyder, D. H., Valsamis, M. P. & Stone, S. H. *Lab. Invest.* **31**, 369-380 (1974).
3. Raine, C. S. in *Multiple Sclerosis—Pathology, Diagnosis and Management* (eds Hallpike, J., Adams, C. W. M. & Tourtellote, W. W.) 413-460 (Chapman & Hall, London, 1983).
4. Lassman, H. K. & Wisniewski, H. M. *Archiv. Neurol.* **36**, 490-497 (1979).
5. Zamvil, S. et al. *Nature* **317**, 355-358 (1985).
6. Zamvil, S. et al. *J. exp. Med.* **162**, 2107-2124 (1985).
7. Petinelli, C. B. & McFarlin, D. E. *J. Immun.* **127**, 1420-1423 (1981).
8. Brostoff, S. W. & Mason, D. W. *J. Immun.* **133**, 1938-1942 (1984).
9. Waldor, et al. *Science* **227**, 415-417 (1985).
10. Petinelli, C. B., Fritz, R. B., Jen-Chou, C. H. & McFarlin, D. E. *J. Immun.* **129**, 1209-1214 (1982).
11. Fritz, R. B., Jen-Chou, C. H. & McFarlin, D. E., *J. Immun.* **130**, 191-194 (1983).
12. Fritz, R. B., Skeen, M. J., Jen-Chou, C. H., Garcia, J. & Egorov, I. K., *J. Immun.* **134**, 2328-2332 (1985).
13. Schwartz, R. H., Fox, B. S., Frago, E., Chen, C. & Singh, B. *J. Immun.* **135**, 2598-2608 (1985).
14. Haskins, K. et al. *J. exp. Med.* **157**, 1149-1169 (1983).
15. Kaye, J., Porcellii, S., Tite, J., Jones, B. & Janeway, C. A. Jr. *J. exp. Med.* **158**, 836-856 (1983).
16. Martenson, R. E. *Prog. clin. biol. Res.* **146**, 511-521 (1984).
17. Eylar, E. H., Caccam, J., Jackson, J. J., Westall, F. C. & Robertson, R. B. *Science* **168**, 1220-1223 (1970).
18. McFarlin, D. E., Blank, S. E., Kibler, R. F., McKneally, S. & Shapira, R. *Science* **179**, 478480 (1972).
19. Teitelbaum, D. C., Webb, C., Arnon, R. & Selà, M. *Cell. Immun.* **29**, 265-271 (1977).

20. Kibler, R. F. et al. *J. exp. Med.* **146**, 1323-1331 (1977).
21. Hashim, G. A., Carvalho, E. F. & Shapiro, R. D. *J. Immun.* **121**, 665-670 (1978).
22. Arnon, R. *Immun. Rev.* **55**, 5-30 (1981).
23. Yanagi, Y. et al. *Nature* **308**, 145-149 (1984).
24. Hedrick, S. M., Cohen, D. I., Nielsen, E. A. & Davis, M. M. *Nature* **308**, 149-153 (1984).
25. Chien, Y. et al. *Nature* **312**, 31-35 (1984).
26. Yague, J. et al. *Cell* **42**, 81-87 (1985).
27. Fink, P. J., Matis, L. A., McEllingott, D. L., Bookman, M. & Hendrick, S. M. *Nature* **321**, 219-226 (1986).
28. Goverman, J. et al. *Cell* **40**, 859-867 (1985).
29. Erickson, B. W. & Merrifield, R. B. in *The Proteins* Vol. II (ed. Neurath, H.) 255-527 (Academic, New York, 1976).
30. Rothbard, J. B., Fernandez, R. & Schoolnick, G. K. *J. exp. Med.* **160**, 208-221 (1984).
31. Kimoto, M. & Fathman, C. G. *J. exp. Med.* **152**, 759-770 (1980).
32. Oj, V. T., Jones, P. P., Goding, J. W. & Herzenberg, L. A. *Clin. Topics Microbiol. Immun.* **81**, 115-129 (1978).
33. Ozato, K., Mayer, N. & Sachs, D. H. *J. Immun.* **124**, 533-540 (1980).

Interaction of peptide antigens and class II major histocompatibility complex antigens

Jean-Gerrard Guillet, Ming-Zong Lai, Thomas J. Briner, John A. Smith* & Malcolm L. Gelfert

Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

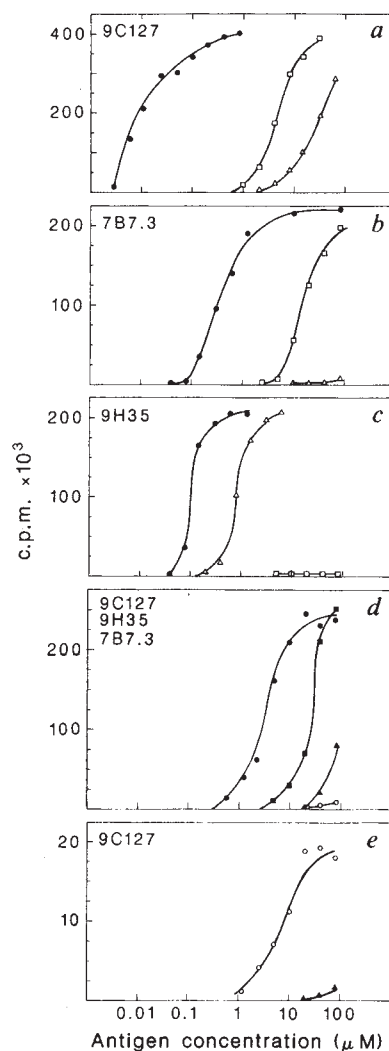
* Departments of Molecular Biology and Pathology, Massachusetts General Hospital and Department of Pathology, Harvard Medical School, Boston, Massachusetts 02114, USA

T lymphocytes require a foreign antigen to be presented on a cell surface in association with a self-transplantation antigen before they can recognize it effectively. This phenomenon is known as major histocompatibility complex (MHC) restriction¹. It is not clear how an incalculably large number of foreign proteins form unique complexes with a very limited number of MHC molecules. We studied the recognition properties of T cells specific for a peptide derived from bacteriophage λ cI protein. Analogues of this peptide, as well as peptides derived from other unrelated antigens which can be presented in the context of the same MHC molecule, can competitively inhibit activation of these T cells by the cI peptide. Furthermore, these unrelated antigens can stimulate cI-specific T cells if certain specific amino-acid residues are replaced. Here we suggest a model in which all antigens give rise to peptides that can bind to the same site on the MHC molecule. T-cell recognition of this site (which is presumed to be polymorphic) with or without antigen bound can explain self-selection in the thymus and MHC restriction.

We immunized BALB/c mice with the 102 amino-acid N-terminal domain of bacteriophage λ cI protein (cI) and produced cI-specific, class II-restricted T-cell hybridomas. More than 75% of specific hybridomas could be stimulated by a peptide comprised of residues 12-26 (P12-26) in the context of 1-A^d. Independently isolated T cells could be divided into at least three groups on the basis of their reactivity to truncated analogues of P12-26 (Fig. 1). Following Rock and Benacerraf² and others³, who showed competition for presentation between related antigens, we examined whether inactive truncated analogues could inhibit activation of T cells stimulated with the parent peptide or active truncated analogues. In the case of the hybridoma 7B7.3, P12-24 can inhibit activation by P15-26, but not in the case of hybridoma 9H35 (see Fig. 2a). Failure to obtain competition has also been noted in other systems^{1,4}, suggesting that competition depends on the T cell involved. It may well be that the active peptide and the analogues compete for binding to the same site on I-A (ref. 5) but the T cell enhances the avidity of the active peptide for I-A so that competition cannot be seen. This hypothesis agrees with recent findings from several laboratories^{6,7}. To test the hypothesis that other antigens which can be presented in the context of 1-A^d inhibit cI peptide stimulation, we screened 14 overlapping peptides derived from

Fig. 1 *a-c*, Response of cI-specific T-cell hybridomas to various peptides derived from bacteriophage cI protein. Activity was measured with P12-26 (residues 12-26) (●), P15-26 (residues 15-26) (□) and P12-24 (residues 12-24) (△) as measured by IL-2 production. *a*, Hybridoma 9C127; *b*, hybridoma 7Bn,3; *c*, hybridoma 9H35. Values are the arithmetic mean of triplicate samples. *d, e*, The stimulation of cI-specific T cells by analogues of peptides derived from *S. aureus* nuclease and chicken ovalbumin. *d*, Hybridomas 9C127 (●), 9H35 (▲) and 7B7.3 (■) were incubated in the presence of [Lys 73, Ile 75] Nase I (see Fig. 3) at the concentration indicated and activation was measured by IL-2 release as described below. The activity in the presence of Nase I (see Fig. 3) for all the T-cell hybridomas is also shown (○). *e*, Hybridoma 9C127 was incubated in the presence of [Lys 331, Ile 333, Lys 336] Ova I (○) (see Fig. 3) and Ova I (▲) (see Fig. 3), as described above.

Methods. The T-cell hybridomas were derived from BALB/c mice immunized with bacteriophage λ repressor cI protein N-terminal fragment (residues 1-102), as described⁴. Those T cells reacting to a synthetic peptide (residues 12-26) were selected. The complete sequence of λ repressor cI protein is reported by Sauer and Andereg¹³. All peptides used in this letter were made by the solid phase method of Merrifield¹⁴ using an automated Applied Biosystems 430A peptide synthesizer. The peptides were desalted on a G-25 column (2.5 × 90 cm) equilibrated with 4 M glacial acetic acid and then lyophilized twice after resuspension in distilled water. 5×10^4 A20 antigen presenting cells were incubated in 200 μ l in 96-well Costar plates with 5×10^4 hybridoma T cells at the indicated concentration of peptide. After 24 h of culture, supernatant (50 μ l) was harvested and assayed for IL-2 concentration by the following incorporation of 3 H-thymidine into the IL-2 dependent CTL-L cell line (10^4 cells per well). Twenty-four hours after culture of T-cell supernatant in the presence of CTL-L cells, 1 μ Ci of 3 H-thymidine was added per well. Six hours later the cells were harvested by an automated cell harvester (Skatron Inc.) and thymidine incorporation was measured by scintillation counting.



the antigen *Staphylococcus aureus* nuclease (Nase)⁸. One peptide (residues 61-80) was a potent inhibitor of all three cI-specific T cells tested (Fig. 2b). Note that the stimulating analogue showing the lowest apparent avidity was used for each T cell respectively (see Fig. 2, legend) to allow observation of inhibition.

If the inhibition observed with the Nase peptide is biologically relevant, that is, if its ability to bind to a specific site on the I-A molecule is related to its ability to be presented, then Nase peptide P61-80 should also be the target peptide for Nase reactivity in the H-2^d mouse. This is indeed the case⁸. As a control, the Nase peptide P81-100 is the target peptide for activation in the context of I-E^k and does not act as an inhibitor in this system⁸. Examination of the sequence of residues of 61-80 revealed that residues 66-78 (Nase I) were remarkably similar to the cI peptide P12-24 (Fig. 3). There are only two residues (positions 73 and 75) that differ significantly; all other changes are conservative. If Nase I binds to exactly the same site as P12-24 and if the structures of the bound peptides are the same, we reasoned that replacement of the residues at positions 73 and 75 of Nase I with the residues found in the equivalent positions on P12-24 might produce a peptide able to stimulate the cI-specific T cells. This is indeed the case (Fig. 1d). Both peptides are equally active for the T cell 9C127, but the T cell 9H35 requires ~100-fold higher concentration of the Nase analogue. Surprisingly, the peptide [Lys 73, Ile 75] Nase I stimulates the T cell 7B7.3, whereas the cI derived peptide (the immunogen analogue) P12-24 does not. The unmodified Nase peptide does not stimulate any cI-specific

T cells.

This suggested to us that all peptides restricted by I-A^d might bind to exactly the same site on I-A and act as inhibitors of cI-specific T cell activation. The peptide derived from ovalbumin, residues 324-336 (Ova I) is I-A^d restricted, and its behaviour has been well characterized^{4,9}. As shown in Fig. 2c, this peptide inhibits all cI-specific T cells tested. However, this peptide has little obvious similarity to P12-24 (Fig. 3). We have shown independently¹⁰ that P12-26 competes with Ova P323-339 for binding to the I-A molecule. We assume that Ova I is capable of forming a structure very similar to P12-24 either in solution or, more likely, once bound to the I-A molecule. If we assume that both peptides bind to the same site on I-A and that the structural requirements for T-cell activation are the same for all antigens, then we should be able to make Ova I activate cI-specific T cells by replacing a few critical residues on the Ova I peptide. Watts *et al.*⁹ have shown that the histidine residue (boxed in Fig. 3) is a critical 'epitope' for Ova-specific T-cell activation. Using this residue to align Ova I with P12-24, we made the peptide [Lys 331, Ile 333, Lys 336] Ova I (see Fig. 3). This peptide can stimulate the cI-specific T-cell hybridoma 9C127 (Fig. 1e), which has the highest affinity for 12-24. The concentration of the Ova analogue peptide required to give half-maximal stimulation of 9C127 is the same as P12-24. This probably reflects its intrinsic affinity for I-A (also as measured in the inhibition assay). The maximal level of stimulation for the 9C127 T cell is 15-20% of that seen with P12-24; this probably reflects a weak affinity of the T cell for the bound antigen. It is plausible that maximal T-cell stimulation requires

Fig. 2 Inhibition of T-cell activation by various peptides. The extent of inhibition was analysed by comparing the amount of IL-2 production in the presence of a fixed concentration of inhibitor and varying the amount of stimulating peptide over a 50-fold range. Inhibition was clearly competitive in that the extent of inhibition seen was greatest at high ratios of inhibitor to activator and least when the ratios were reversed. For simplicity, we present here the inhibition data obtained at a concentration of stimulator which gives 50% of maximal stimulation; the inhibitor concentrations are given below. *a*, The activity of T-cell hybridoma 7B7.3 was measured in the presence of P15-26 (20 μ M) alone as described in the legend to Fig. 1 or in the presence of both P15-26 (20 μ M) and P12-24 (40 μ M). The response of T-cell hybridoma 9H35 was measured in the presence of P12-24 (1 μ M) alone or in combination with P15-26 (60 μ M). The difference in activity seen when two peptides were included in assay cultures compared to the activity seen when stimulator peptide alone was given is presented as the per cent inhibition. *b*, The activity of the T-cell hybridoma 9C127 was measured in the presence of P12-24 (20 μ M) alone, as described in Fig. 1 (legend) or in the presence of both P12-24 (20 μ M) and a synthetic peptide (residues 61-80) derived from the protein *S. aureus* nuclease (Nase I) (60 μ M). The response of the hybridoma 9H35 was measured in the presence of P12-24 (1 μ M) alone or in combination with Nase I (60 μ M). The response of the hybridoma 7B7.3 was measured in the presence of P15-26 (20 μ M) alone or in combination with Nase I (60 μ M) or Nase II (residues 81-100) (100 μ M). The activity of T-cell hybridoma 8I was measured in the presence of P15-26 (0.25 μ M) alone or in combination with Nase I (60 μ M). T-cell hybridoma 8I was derived from immune A/J mice and recognized P15-26 in the context of I-E^k. Accordingly the cell line TA3 was used for antigen presentation. *c*, The same conditions as *b*, except Ova I (25 μ M) was used as an inhibitor.

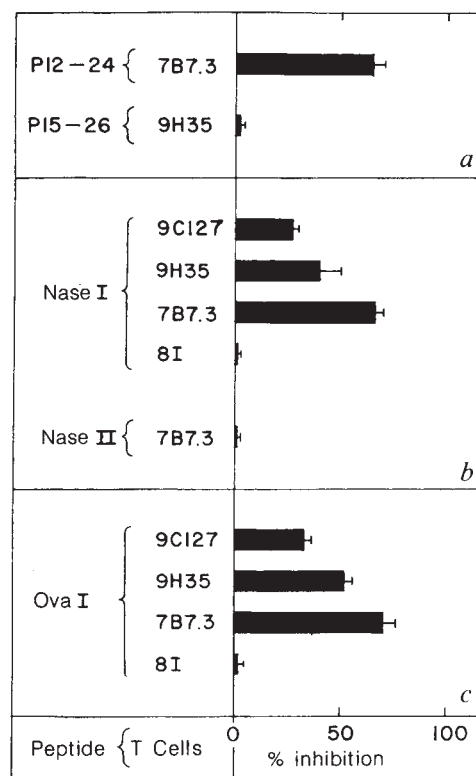


Fig. 3 The wild-type sequence of bacteriophage λ repressor protein residues 12-24 (P12-24), is shown. Nase I represents residues 66-78 from *S. aureus* nuclease and Ova I represents residues 324-336 from chicken ovalbumin. The residues in boxes indicate the locations where modifications to the parent peptides were made. All peptides and their analogues were synthesized and purified as described in Fig. 1. The peptide [Lys 73, Ile 75] Nase I has the amino-acid sequence of Nase I, except that the residues at positions 73 and 75 were replaced by the residues present at the boxed positions of P12-24. The peptide [Lys 331, Ile 333, Lys 336] Ova I has the amino-acid sequence of Ova I, except that the residues at positions 331, 333 and 336 were replaced by the residues at the boxed positions of P12-24.

P12-24	LEU - GLU - ASP - ALA - ARG - ARG - LEU - LYS - ALA - ILE - TYR - GLU - LYS
Nase I	VAL - GLU - ASN - ALA - LYS - LYS - ILE - GLU - VAL - GLU - PHE - ASP - LYS
Ova I	SRR - GLN - ALA - VAL - HIS - ALA - ALA - HIS - ALA - GLU - ILE - ASN - GLU

more than the two identical amino-acid contacts provided by the analogue.

These results suggest that the class II transplantation antigen is a simple receptor containing a single binding site for peptides able to form similar general structures once bound. The function of the transplantation antigen would be to position the antigen for T-cell binding uniquely. It has been postulated that all T-cell antigens are peptides capable of forming amphipathic α helices¹¹; it is not likely that the peptide P12-26 described here will do so. However, it is attractive to consider that a structure like an α helix or β sheet is the common structure that immunogenic peptides form either in solution, or, more likely, when bound to the transplantation antigen. Given recent results¹², it is also reasonable to propose that similar models can be constructed for class I-restricted antigen presentation. The 'docking' of the T-cell receptor onto the transplantation antigen which allows it to interact with the antigen uniquely requires an intrinsic affinity of the T-cell receptor for the area of the protein surrounding the peptide binding site. This area is the same for all antigens presented on a given transplantation antigen molecule, and is presumably the site learned as 'self' in the thymus. It follows that the binding site and its local area is the polymorphic region of the I-A molecule which imparts unique qualities to both the peptide binding site itself (controlling which amino-acid sequences bind best to the site) and the

self recognition site(s). We envisage that the transplantation antigen has one peptide binding site. The properties of the site (α helical-binding) would be imposed by the general structure of the transplantation antigen (class I or class II).

This model provides a framework for the construction of synthetic immune agonists and antagonists. Further physical and chemical data are required to test these proposals.

This work is supported by grant AI13357 from NIH to M.L.G. and by a grant from Hoechst Aktiengesellschaft (FRG) to J.A.S. M.Z.L. is supported by a postdoctoral fellowship, DRG-853, from the Damon Runyon-Walter Winchell Cancer Fund.

Received 2 July; accepted 11 September 1986.

- Schwartz, R. H. *A. Rev. Immun.* **3**, 237-261 (1985).
- Rock, K. L. & Benacerraf, B. *J. exp. Med.* **157**, 1618-1634 (1983).
- Werdlein, O. *J. Immun.* **129**, 1883-1891 (1982).
- Shimonkevitz, R., Colon, S., Kappler, J. W., Marrack, P. & Grey, H. M. *J. Immun.* **133**, 2067-2074 (1984).
- Babbitt, B. P., Allen, P. M., Matsueda, G., Haber, E. & Unanue, E. R. *Nature* **317**, 359-360 (1985).
- Watts, T. H., Gaub, H. E. & McConnell, H. M. *Nature* **320**, 179-181 (1986).
- Ashwell, S. D. & Schwartz, R. H. *Nature* **320**, 176-179 (1986).
- Finnegan, A. *et al. J. exp. Med.* **164**, 899-910 (1986).
- Watts, T. H., Gariepy, J., Schoolnick, G. K. & McConnell, H. M. *Proc. natn. Acad. Sci. U.S.A.* **82**, 5480-5484 (1985).
- Guillet, J.-G. *et al. Proc. natn. Acad. Sci. U.S.A.* (submitted).
- Delisi, C. & Berzofsky, J. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7048-7052 (1985).
- Townsend, A. R. M. *et al. Cell* **44**, 959-968 (1986).
- Sauer, R. T. & Andereg, R. *Biochemistry* **17**, 1092-1100 (1978).
- Merrifield, R. B., *J. Am. chem. Soc.* **85**, 2149-2154 (1963).

Identification of a common class of high affinity receptors for both types of porcine interleukin-1 on connective tissue cells

T. A. Bird & J. Saklatvala

Strangeways Research Laboratory, Worts Causeway, Cambridge CB1 4RN, UK

Interleukin-1 (IL-1) is the name given to the polypeptides produced by activated mononuclear phagocytes which were originally defined as lymphocyte activating factors (LAF)¹. Administration of IL-1 *in vivo* causes fever² and synthesis of acute phase proteins³. *In vitro* they have been shown to cause cartilage and bone resorption⁴⁻⁷, and to stimulate fibroblasts and chondrocytes to make prostaglandins and latent collagenase^{4,8}. IL-1 has therefore been proposed to be an important inflammatory mediator⁹ and may be involved in the destruction of cartilage and bone that is a feature of rheumatoid arthritis and other inflammatory diseases of joints. We therefore looked for IL-1 receptors on connective tissue cells which might be targets for therapeutic intervention. Here we report the iodination, to high specific activity and with retention of full biological potency, of the two types of natural porcine IL-1. These ligands have been used to demonstrate high affinity dissociation constant ($\sim 10^{-10}$ M) specific binding sites on pig chondrocytes and synovial fibroblasts, human dermal fibroblasts and murine osteoblasts (3,000–5,000 sites per cell). Most interestingly, the two different IL-1 proteins show a similar affinity for a common class of receptors.

Two types of IL-1 have been isolated from culture supernatants of porcine mononuclear leucocytes stimulated with Concanavalin A. These are immunologically different proteins that have identical relative molecular masses (M_r s) of 21,000 (21K) on SDS polyacrylamide gel electrophoresis (PAGE) and have been characterized elsewhere⁵. One is acidic, having an isoelectric point (pI) of 5 (IL-1/5) and the other basic (pI of 8). These proteins are probably the porcine counterparts of the human IL-1 α (which has a pI of 5) and the human IL-1 β (which has a pI of 7 (ref. 10)) since an antiserum to porcine IL-1/5 cross reacts with human IL-1 α and the N-terminal amino-acid sequence of porcine IL-1/8 shows 35% homology with that of human IL-1 β (J.S., unpublished observations).

Attempts to label the different types of porcine IL-1 by direct oxidative methods proved unsuccessful, resulting in loss of biological activity. However, we were able to produce 125 I-labelled preparations of IL-1 which were fully active in a cartilage-resorption assay (Fig. 1b) using di- 125 I-Bolton and Hunter reagent. The labelled proteins migrated as single bands upon SDS-PAGE, in positions expected for the native 21K proteins (Fig. 1a). The specific activity in each case was 470–580 Ci mmol⁻¹. 125 I-IL-1 could be stored at -20°C for at least 6 weeks without loss of binding affinity. We initially studied the interaction of 125 I-IL-1/8 and 125 I-IL-1/5 with monolayers of pig synovial fibroblasts. Figure 2a shows the time course for specific (displaceable) binding of 125 I-IL-1/8 (2.4×10^{-10} M) to pig synovial fibroblasts at various temperatures. The rate of binding increased with increasing temperature; equilibrium was only attained after 6 h of incubation at 19°C . To examine whether some 125 I-IL-1/8 was internalized at 37°C parallel cultures were briefly exposed to an acidic glycine buffer¹² to remove surface-bound ligand. 85% of the radioactivity could be removed from monolayers incubated at 4°C or 19°C in this way but, as shown in Fig. 2a, a large proportion of the radioactivity bound at 37°C rapidly became resistant to this treatment and was presumably internalized. The amount of acid-labile radioactivity bound at 37°C reached a steady state within one hour; the fact that a

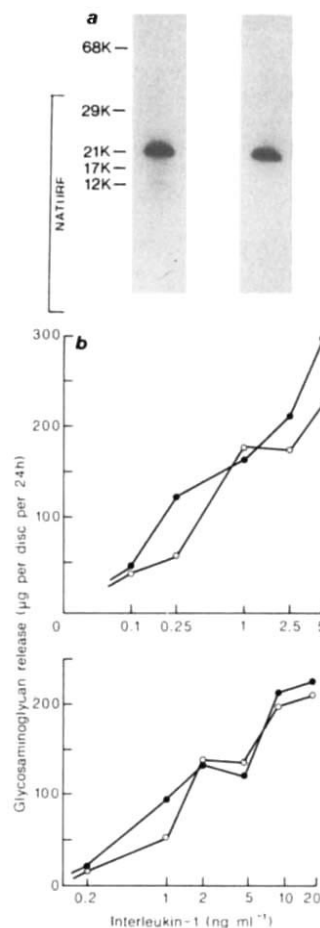


Fig. 1 a, Autoradiographs of 125 I-IL-1/8 (lane 1) and 125 I-IL-1/5 (lane 2) electrophoresed on 12.5% SDS-polyacrylamide slab gels in an ammediol buffer system¹⁶ with reduction. The M_r s of Coomassie brilliant blue-stained standard proteins run in parallel are indicated in kilodaltons; b, comparison of 125 I-IL-1/5 (●; upper panel) and 125 I-IL-1/8 (●; lower panel) with the corresponding native interleukin 1 (○) in a cartilage resorption assay. Vertical axis, glycosaminoglycan release (μg per disc per 24 h); horizontal axis, concentration of IL-1 (ng ml^{-1}).

Methods. Purified IL-1s ($2\text{--}4\text{ }\mu\text{g}$ estimated from A_{280} , assuming absorbance of a 1 mg ml^{-1} solution in a 1 cm lightpath to be unity) were iodinated by the method of Bolton and Hunter¹⁷. The interleukins, freeze-dried from aqueous solution, were dissolved in $20\text{ }\mu\text{l}$ 0.1 M borate buffer ($\text{pH } 8.5$) and added to 1 mCi N -succinimidyl 3-(4-hydroxy-3,5- 125 I-diiodophenyl) propionate freshly dried from a solution in benzene. After agitation of the mixture on ice for 30 min the reaction was quenched by addition of $50\text{ }\mu\text{l}$ 1 M glycine in borate buffer. Iodinated proteins were separated from byproducts by gel filtration on Sephadex G-25 which was equilibrated and eluted with 0.05 M sodium phosphate buffer ($\text{pH } 7.5$) containing 0.25% (w/v) gelatin. Because 125 I-glycine and free 125 I-Bolton and Hunter reagent were quantitatively eluted from the column, residual radioactivity in the gel slurry was counted to estimate retained 125 I-IL-1. Recovery of labelled protein was typically 75% . After correcting for recovery, aliquots of labelled IL-1 were tested against unlabelled material for their ability to stimulate cartilage resorption⁴. Disks (2 mm) of bovine nasal cartilage were maintained in 96-well microtitre plates in $150\text{ }\mu\text{l}$ of Dulbecco's modified Eagles Medium (DMEM) containing 5% (v/v) bovine serum and test samples for 24 h at 37°C in CO_2/air ($1:19$). Chondroitin sulphate released into the medium was measured using the metachromatic dye dimethyl methylene blue. Each point represents the mean of four disks. Release of chondroitin sulphate in the absence of interleukin 1 was less than $30\text{ }\mu\text{g}$ per disk. Note that the different dose/response relationships in Fig. 1b arise from the fact that different batches of cartilage were used for each assay. When the two types of IL-1 are tested on the same batch of cartilage, they have the same specific bioactivity⁵.

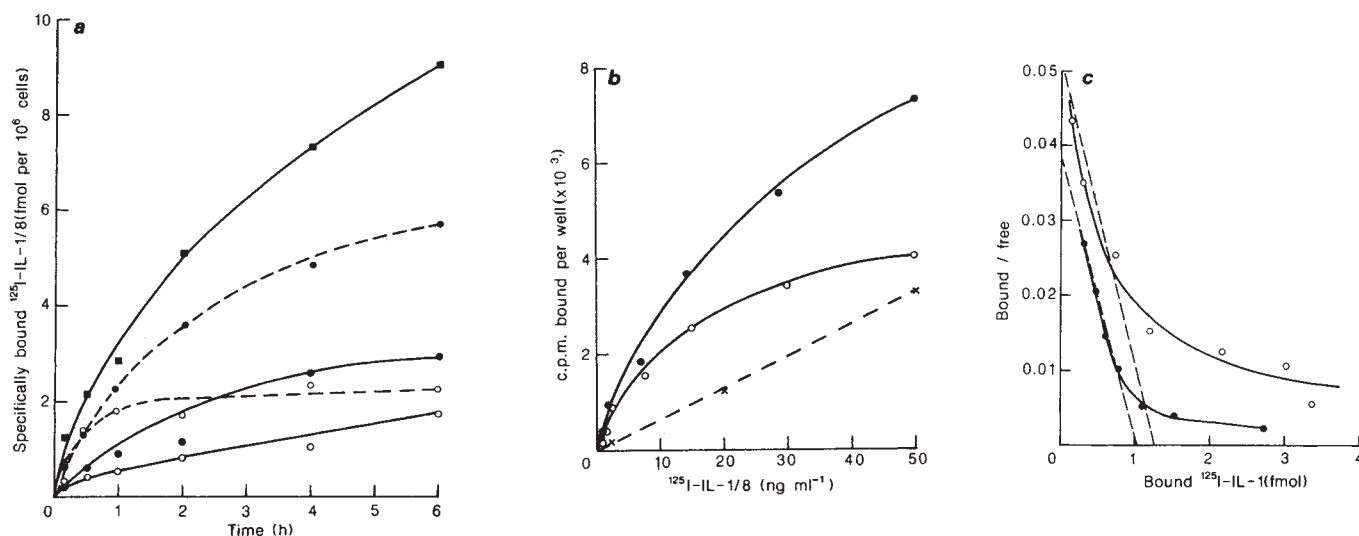


Fig. 2 *a*, Time course of ^{125}I -IL-1/8 binding to porcine synovial fibroblasts (PSF) at various temperatures, and internalization of ligand at 37°C. $\blacksquare$ — $\blacksquare$, binding at 37°C; $\bullet$ — $\bullet$, binding at 19°C; $\circ$ — $\circ$, binding at 4°C; $\circ$ — $\circ$, acid-labile radioactivity; $\bullet$ — $\bullet$, residual, SDS soluble radioactivity; *b*, saturable binding of ^{125}I -IL-1/8 to PSF incubated for 4 h at 19°C; $\bullet$, total binding; $\circ$, specific binding; $\times$, nonspecific binding. *c*, Scatchard analysis of ^{125}I -IL-1/5 ($\bullet$) and ^{125}I -IL-1/8 ($\circ$) to PSF at 19°C.

Methods. Monolayers of PSF, isolated from metacarpophalangeal joints of freshly-killed young pigs by enzymatic disaggregation¹⁸ were maintained in DMEM plus 10% (v/v) heat-inactivated fetal calf serum until required. Newly confluent monolayers (1.2×10^5 cells) in 4 cm² plastic wells were washed twice with Hank's balanced salts solution supplemented with 0.1% BSA and 20 mM HEPES pH 7.4 (binding buffer), then incubated on a rotatory platform (70 cycles min⁻¹) for the indicated times at 4°C, 19°C or 37°C with 400 μl per well binding buffer containing 2.4×10^{-10} M ^{125}I -IL-1/8. Monolayers were then rinsed with 3 changes of ice cold binding buffer (1 ml), solubilized with 0.5 ml 1% SDS and counted in a gamma counter. In some cases, monolayers incubated with ^{125}I -IL-1 at 37°C were rinsed with binding buffer and kept on ice for 15 min in 1 ml of 50 mM glycine/150 mM NaCl pH 3.0 before solubilization in 1% SDS. The acid-labile radioactivity and residual, SDS soluble radioactivity were counted separately. Each point is the mean of three determinations from which nonspecific binding (measured in the presence of 5×10^{-8} M partially purified IL-1) has been subtracted. Conditions in *b* and *c* were the same with varying ^{125}I -IL-1/8 concentrations at 19°C for 4 h. Each point is calculated from the average values of four determinations. Straight lines were fitted, where appropriate, by the method of least squares.

pool of surface-bound ligand was maintained despite the appearance of radioactivity in an intracellular compartment suggests receptor recycling such as that described for the LDL receptor¹³. For subsequent studies, cells were incubated with labelled IL-1 at 19°C for 4 hours. Figure 2*b* shows that binding of ^{125}I -IL-1/8 to pig synovial fibroblasts at 19°C was saturable, with half maximal binding occurring at 4×10^{-10} M (9 ng ml⁻¹); nonspecific binding was 10–30% of total binding. Similar data were obtained using ^{125}I -IL-1/5. Scatchard analysis of binding to pig synovial fibroblasts (Fig. 2*c*) gave curvilinear plots resolving into a high affinity component and, usually, a much lower affinity component. Straight lines could be fitted to points lying on the first part of the plots to yield, by extrapolation, capacities of approximately 5,400 and 4,500 sites per cell for ^{125}I -IL-1/8 and ^{125}I -IL-1/5 respectively, with equilibrium dissociation constants (K_D of 1.5×10^{-10} and 1.7×10^{-10} M). High affinity binding was reproducible over a number of experiments using different batches of cells and different preparations of ^{125}I -IL-1, but the contribution of low affinity binding was less consistent. Given the sensitivity of Scatchard analysis to various sources of error at high levels of bound ligand¹⁴, this component may be artefactual and result from imprecise measurement of non-specific binding. Similar experiments were performed using human dermal fibroblasts and porcine articular chondrocytes. In each case linear Scatchard plots were obtained; human fibroblasts bind 4,800 sites per cell with a K_D of 6.8×10^{-11} M; porcine chondrocytes bind 7,000 sites per cell with a K_D of 2.5×10^{-11} M. Reversibility of binding was assessed after preincubation of pig synovial fibroblasts with 10 ng per ml ^{125}I -IL-1/8 (with or without excess unlabelled IL-1). The cells were washed and incubated at 19°C in 2.5 ml per well of binding buffer. When the preincubation was performed at 19°C, 33% of specifically bound ligand was released over 6 hours. The dissociation curve followed a first order model with $K_{-1} = 2 \times 10^5 \text{ s}^{-1}$ and $t = 637$ min. The dissociable radioactivity was TCA-precipitable

(not shown): 55% of the radioactivity bound at 37°C became dissociated over 6 hours of which 40% was TCA-precipitable, indicating that some degradation had occurred.

We wished to test whether both types of ^{125}I -IL-1 interacted with the same receptor sites; Fig. 3*a* shows the results of experi-

Table 1 Single-point equilibrium analysis of binding of ^{125}I -IL-1/5 and ^{125}I -IL-1/8 to various cells

Cell type	Derivation	Species	Molecules bound per cell IL-1/8	IL-1/5
HuGi	Gingival fibroblast	Human	ND	715 $\pm$ 45
Flow 7000	Foreskin fibroblast	Human	3,709 $\pm$ 287	3,429 $\pm$ 120
PSF	Synovial fibroblast	Pig	3,709 $\pm$ 223	2,546 $\pm$ 104
PCC	Chondrocyte	Pig	4,050 $\pm$ 270	ND
PMO	Osteoblast	Mouse	2,605 $\pm$ 800	3,248 $\pm$ 122
3T3	Fibroblast	BALB/c Mouse	ND	2,752 $\pm$ 146
HeLa	Cervical carcinoma	Human	ND	<100
HT1080	Fibrosarcoma	Human	<100	<100
MG63	Osteogenic sarcoma	Human	929 $\pm$ 195	909 $\pm$ 56
UMR 106	Osteogenic sarcoma	Rat	543 $\pm$ 59	ND
M-3	Melanoma Cloudman	Mouse	<100	<100

Chondrocytes were isolated by sequential enzyme digestion¹⁹ of pig metacarpal head cartilage and used in primary culture. Pig synovial fibroblasts, prepared as described in Fig. 1 (legend), were used between passages three and nine. Human gingival fibroblasts, a gift from S. J. Atkinson, were subcultured through six passages before use. Human foreskin fibroblasts (Flow 7000) were used between passages 17 and 21. Neonatal mouse calvarial osteoblasts, prepared by a sequential digestion procedure²⁰ were obtained from Dr J. Heath and used in primary culture. The UMR-106 cell line was obtained from Professor T. Chambers; all other established lines were from Flow Laboratories. All cells were maintained in DMEM supplemented with 10% (v/v) heat-inactivated FCS, and were subcultured, where appropriate, using trypsin-EDTA. Confluent monolayers, in 4 cm² wells, were incubated for 4 h at 19°C with ^{125}I -IL-1/8 (3.57×10^{-10} M) or ^{125}I -IL-1/5 (2.38×10^{-10} M). Specifically bound radioactivity was measured as described for Fig. 2. Cells in additional wells were counted in a haemocytometer after trypsin-EDTA treatment. Results are the mean $\pm$ s.e.m. of four determinations; ND, not determined.

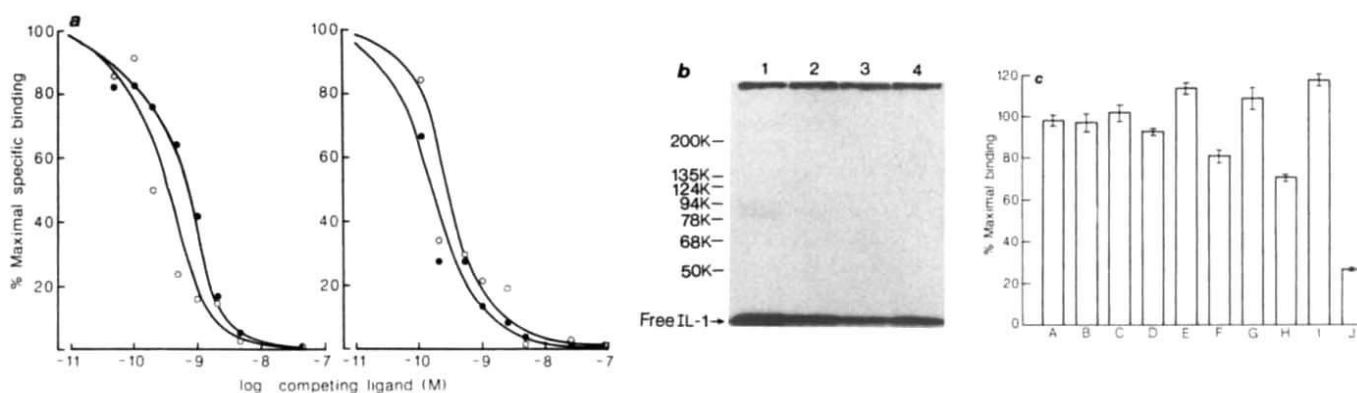


Fig. 3 *a*, Competition of IL-1/5 (○) and IL-1/8 (●) with: left panel, ¹²⁵I-IL-1/5; right panel, ¹²⁵I-IL-1/8 for binding to porcine synovial fibroblasts. Vertical axis, per cent maximal specific binding; horizontal axis, log₁₀ concentration of competing ligand. *b*, Covalent crosslinking of ¹²⁵I-IL-1/8 to intact monolayers. *c*, Specificity of the IL-1 receptor; the percentage of maximum ¹²⁵I-IL-1 binding observed in the presence of various other unlabeled polypeptide hormones and growth factors is shown.

Methods. Experimental conditions in *a* were as for Fig. 2*b* with the addition of the indicated concentrations of unlabelled IL-1; labelled IL-1 was added at 1.2×10^{-10} M. Results are expressed as specific binding relative to binding in the absence of competing ligand and are the means of triplicate determinations. *b*, Monolayers of 3T3 fibroblasts (clone A41; Flow Laboratories) in 25 cm² flasks were incubated for 4 h at 19 °C with 5 ml binding buffer supplemented with 9.5×10^{-10} M ¹²⁵I-IL-1/8 and unlabelled IL-1/5 or IL-1/8 as indicated below. Unbound radioactivity was removed by washing the monolayers with three changes of ice-cold protein-free binding buffer. The flasks each received 5 ml of this buffer followed by 200 µl ethylene glycol bis (succinimidyl succinate) (Pierce) freshly dissolved in dimethyl sulphoxide; final crosslinker concentration was 0.5 mM. After standing the flasks on ice for 45 min, the crosslinker was removed and replaced with 5 ml of a quenching solution (10 mM TrisHCl pH 7.5, 1 mM EDTA). The cells were scraped into this buffer, pelleted by centrifugation and dissolved in 300 µl of 1% SDS containing the protease inhibitors PMSF (10 mM) EDTA (2 mM) and *N*-ethylmaleimide (10 mM). Samples containing equal amounts of radioactivity were adjusted to 125 mM Tris HCl pH 6.8, 2% SDS, 5% 2-mercaptoethanol, boiled for 5 min and electrophoresed on SDS-polyacrylamide (5% stacking gel, 7.5% resolving gel)²¹. The gel was dried under vacuum and exposed to X-ray film for 7 days at -80 °C. Track, 1, no competing ligand; track 2, porcine IL-1/8 (1 µg ml⁻¹); track 3, porcine IL-1/5 (1 µg ml⁻¹); track 4, recombinant murine IL-1 (1 µg ml⁻¹). The positions and *M_s* (in kilodaltons) of α₂-macroglobulin, chick skin collagen, phosphorylase a, transferrin, serum albumin and goat IgG are given. *c*, Porcine synovial fibroblasts were incubated with 2.4×10^{-10} M ¹²⁵I-IL-1/5 and various unlabeled polypeptide hormones and growth factors. A, synthetic human angiotensin II (2 µg ml⁻¹); B, human vasopressin (2 µg ml⁻¹); C, bovine parathyroid hormone (2 µg ml⁻¹); D, salmon calcitonin (2 µg ml⁻¹); E, mouse submaxillary gland epidermal growth factor (2 µg ml⁻¹); F, human platelet-derived growth factor (1 µg ml⁻¹); G, bovine insulin (3 µg ml⁻¹); H, mouse submaxillary gland 2.5 S nerve growth factor (90% pure, containing some renin-like activity 2 µg ml⁻¹); I, recombinant human tumour necrosis factor (2 µg ml⁻¹); J, porcine IL-1/5 (1 µg ml⁻¹). Values are expressed as the means ± s.e.m. of radioactivity bound in triplicate determinations relative to an assay without added competitor. Samples A, B, and G were from Sigma; E and F were from Bethesda Research Laboratories; C and D were obtained from Dr J. K. Heath, Cambridge; H was a gift from Dr R. Lindsay, Sandoz Institute for Medical Research, London; and sample I was provided by Dr. M. Palladino of Genentech, San Francisco.

ments in which unlabelled IL-1 (both types) was used to displace labelled IL-1 (1.2×10^{-10} M) from pig synovial fibroblasts. It is clear that the two different molecules bind to a common class of receptor sites with an approximately equal affinity. Very similar results were obtained using human dermal fibroblasts and porcine articular chondrocytes (not shown). Additional evidence that both types of IL-1 interact with a common site on fibroblasts is provided by crosslinking studies on pig synovial fibroblasts and murine 3T3 fibroblasts. Similar results are obtained with both cell types but because the efficiency of crosslinking on 3T3 is greater, only these results are shown (Fig. 3*b*). When ¹²⁵I-IL-1/8 was bound to fibroblasts in the presence of a homobifunctional crosslinker, faint high molecular weight complexes (95K and 190K) could be seen after SDS-PAGE and autoradiography. The patterns were similar when samples were analysed under reducing and non-reducing conditions. Inclusion in the incubation medium of a high concentration of either unlabelled porcine IL-1 or murine recombinant IL-1 prevented the formation of these complexes indicating that they are involved in the specific binding of IL-1 of either type.

We examined the specificity of the IL-1 receptor further by including a series of polypeptide growth factors and hormones in the binding assay (Fig. 3*c*). Of the ten factors tested, only IL-1 was able to compete effectively for binding of ¹²⁵I-IL-1/5 to pig synovial fibroblasts. We have screened other cells for IL-1 binding (see Table 1); receptor expression varies from a high level in chondrocytes and osteoblasts which respond strongly to IL-1 by producing collagenase and prostaglandin (unpublished results), to cells without measurable binding (HeLa and

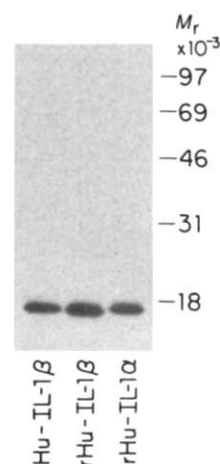
HT1080 which do not produce PGE₂ in response to IL-1 and may be relatively insensitive. Our findings for fibroblasts are quantitatively similar to those recently reported¹⁵ for binding of human ¹²⁵I-IL-1 (p17) which was 95% inactivated by Bolton-Hunter labelling fibroblasts to human and mouse fibroblasts, and they substantiate an initial report²² of cross competition for binding of recombinant human IL-1 on murine T cells. The present report indicates the suitability of the different types of porcine IL-1 as probes for receptors, as they can be radiolabelled preserving full activity, and provides the first evidence that the two forms show similar affinity for a common receptor on relevant target cells.

We thank Dr M. Palladino for the sample of recombinant TNF, Dr P. Lomedico for the gift of recombinant murine IL-1 and the Arthritis and Rheumatism Council for their generous support.

Received 29 January; accepted 10 September 1986.

1. Mizel, S. B., *Immun. Rev.* **63**, 51-72 (1982).
2. Murphy, P. A., Simon, P. L. R. & Willoughby, W. F. *J. Immun.* **124**, 2498-2501 (1980).
3. Sztein, M. B. *et al. Cell Immun.* **63**, 164-176 (1981).
4. Saklatvala, J., Pilsworth, L. M. C., Sarsfield, S. J., Gavrilovic, J. & Heath, J. K. *Biochem. J.* **224**, 461-466 (1984).
5. Saklatvala, J., Sarsfield, S. J. & Townsend, V. *J. exp. Med.* **162**, 1208-1222 (1985).
6. Heath, J. K., Saklatvala, J., Meikle, M. C., Atkinson, S. J. & Reynolds, J. J., *Calc. Tissue Int.* **37**, 95-97 (1985).
7. Gowen, M., Wood, D. D., Ihrie, E. J., McGuire, M. K. B. & Russell, R. G. G. *Nature* **306**, 378-380 (1983).
8. Mizel, S. B., Dayer, J.-M., Krane, S. M. & Mergenhagen, S. E., *Proc. natn. Acad. Sci. U.S.A.* **78**, 2474-2477 (1981).
9. Oppenheim, J. J. & Gery, I. *Immun. Today* **3**, 113-119 (1982).
10. March, C. J. *et al. Nature* **315**, 641-647 (1985).

11. Saklatvala, J., Sarsfield, S. J. & Wood, D. D. *Br. J. Rheumat.* **24**, (Suppl) 68-71 (1985).
12. Ascoli, M., *J. Biol. Chem.* **257**, 13306-13311 (1982).
13. Goldstein, J. L. & Brown, M. S. *J. cell Sci. Suppl.* **3**, 131-137 (1985).
14. Mendel, C. M. & Mendel, D. B., *Biochem. J.* **228**, 269-272 (1985).
15. Dower, S. K., *et al. J. exp. Med.* **162**, 501-515 (1985).
16. Wyckoff, M., Rodbard, D. & Chrambach, A. *Analyt. Biochem.* **78**, 459-482 (1977).
17. Bolton, A. E. & Hunter, W. M. *Biochem. J.* **133**, 529-539 (1973).
18. Pilsworth, L. M. C. & Saklatvala, J. *Biochem. J.* **216**, 481-489 (1983).
19. Schwartz, E. R. in *Skeletal Research* (eds Simmons, D. J. & Kunin, A. S.) 61-80 (Academic, New York, 1979).
20. Heath, J. K., Atkinson, S. J., Meikle, M. C. & Reynolds, J. J. *Biochim. biophys. Acta* **802**, 151-154 (1984).
21. Laemmli, U. K. *Nature* **227**, 680-685 (1970).
22. Dower, S. K., Call, S. M., Gillis, S. & Urdal, D. L. *Proc. natn. Acad. Sci. U.S.A.* **83**, 1060-1064 (1986).



The cell surface receptors for interleukin-1 α and interleukin-1 β are identical

Steven K. Dower, Shirley R. Kronheim, Thomas P. Hopp, Michael Cantrell, Michael Deeley, Steven Gillis, Christopher S. Henney & David L. Urdal

Immunex Corporation, 51 University Street, Seattle, Washington 98101, USA

Interleukin-1 (IL-1) is a factor that can induce proliferation of murine T lymphocytes^{1,2} and can elicit a variety of other biological responses. These include bone resorption³, fibroblast proliferation⁴, acute phase protein release from hepatocytes⁵, cartilage breakdown⁶ and fever⁷. This spectrum of activities is consistent with a role for IL-1 as a mediator of inflammation. Recently, sequence data have shown that there are at least two members of the IL-1 family; these distantly related proteins have been termed IL-1 α and IL-1 β ⁸⁻¹¹. We have found previously that both murine T cells and fibroblasts possess a relative molecular mass (M_r) ~80,000 (80K) plasma membrane receptor for human IL-1 β ^{12,13}. We show here that the receptor for IL-1 α on both murine and human cells is identical to that for IL-1 β . This result raises the issue of what separation, if any, there might be between the biological activities of IL-1 α and IL-1 β .

Both IL-1 α and IL-1 β are encoded by complementary DNAs for proteins of ~30K⁸⁻¹¹, whereas the active forms found in macrophage culture supernatants are ~17K^{3,14-16}. The start positions within the precursors for murine IL-1 and human IL-1 β (nIL-1 β) have been established by N-terminal protein sequence analysis^{8,10}. Recombinant human IL-1 α (rIL-1 α) and IL-1 β (rIL-1 β) were produced in *Escherichia coli* from truncated cDNA clones encoding polypeptides identical to the secreted forms, and subsequently purified to homogeneity (S.R.K., unpublished). nIL-1 β was purified from culture supernatants and labelled with di-iodo (¹²⁵I) Bolton-Hunter reagent as described previously^{12,15}. rIL-1 β was labelled with Bolton-Hunter reagent, and rIL-1 α was labelled with sodium (¹²⁵I) iodide using a modified chloramine-T method¹⁷. For all three lymphokine preparations, radioiodination yielded specific activities in the range 1×10^{15} to 5×10^{15} c.p.m. mmol⁻¹ (0.4-2 atoms I per molecule protein) and SDS-polyacrylamide gel electrophoresis (PAGE) revealed a single labelled polypeptide of 17.5K (Fig. 1), consistent with previously reported^{13,10,14-16} values for IL-1.

All these labelled preparations were >98% TCA precipitable, thus the ¹²⁵I was covalently bound to protein. Comparison of biological activities on murine T lymphoma cells¹⁸ between labelled and unlabelled IL-1 showed that rIL-1 β , like nIL-1 β ¹², suffered some loss of activity on labelling, but rIL-1 α retained 100% biological activity (1×10^{12} U mg⁻¹) after labelling and >95% could be bound by IL-1 receptor bearing cells.

Table I summarizes the binding of ¹²⁵I-nIL-1 β , ¹²⁵I-rIL-1 β

Fig. 1 SDS-PAGE analysis of radiolabelled IL-1 preparations. ¹²⁵I-IL-1 preparations were boiled for three minutes in sample buffer and analysed on a 10-20% polyacrylamide gradient slab gel as described elsewhere¹². Track 1, human IL-1 β ; track 2, recombinant human IL-1 β ; track 3, recombinant human IL-1 α . **Methods.** IL-1 preparations to be used for radiolabelling were checked for purity and approximate protein concentration by SDS-PAGE and silver staining; the exact concentration was later determined by amino-acid analysis on an LKB 4150 amino-acid analyser and comparison with the predicted composition from the cDNA sequence⁸. nIL-1 β was labelled using di-iodo (¹²⁵I) Bolton-Hunter reagent (NEN) as described previously¹². Briefly 1 μ g (0.06 nmoles) of nIL-1 β in 50 μ l of borate (0.02 M) buffered saline (0.15 M) was added to 10 μ l of sodium borate (0.1 M) buffered saline (0.75 M) pH 8.5; the IL-1 was reacted with 1 mCi Bolton-Hunter reagent according to the manufacturer's instructions for 12 h at 8°C. Subsequently 30 μ l of 2% gelatin and 5 μ l of 1 M glycine ethyl ester were added and the protein separated from unreacted Bolton-Hunter reagent on a 1 ml bed volume Biogel P6 (Biorad) column. Routinely ~5% of the Bolton-Hunter reagent was incorporated into protein. Recombinant IL-1 β was labelled with di-iodo Bolton-Hunter reagent by essentially the same protocol, except that 10 μ g (0.57 nmol) of protein in 10 μ l of phosphate (0.015 M) buffered saline (0.15 M) pH 7.2 were mixed with 10 μ l of sodium borate (0.1 M) buffered saline (0.15 M) pH 8.5 and reacted with 1 mCi (0.23 nmoles) of Bolton-Hunter reagent. Routinely 50-60% incorporation of label was observed. Recombinant IL-1 α was labelled with Na(¹²⁵I) using chloramine-T. Briefly 10 μ g of rIL-1 α (0.57 nmol) in 10 μ l of phosphate buffered saline were added to 2.5 mCi (1.0 nmol) of sodium iodide in 25 μ l of 0.05 M sodium phosphate pH 7.0, reaction was initiated by addition of 30 μ l of 1.4×10^{-4} M chloramine-T (4.2 nmol, Sigma). After 30 min on ice the reaction mixture was fractionated by gel filtration as described above. Routinely 40-50% of ¹²⁵I was incorporated into protein. Precipitation of the ¹²⁵I-IL-1 preparations in 1% gelatin carrier using 10% trichloroacetic acid, or rechromatography on gel filtration columns both showed that >98% of the ¹²⁵I label was protein bound after P6 purification.

and ¹²⁵I-rIL-1 α to LBRM-33-1A5 murine T lymphoma cells, human dermal fibroblasts, BALB-c/3T3 cells and ARH77, a human B lymphoblastoid cell line. The number of sites on the LBRM-33-1A5 cells (Table 1) was at the high end of the range for 25 experiments, in which we observed considerable fluctuation in the number of receptors detected (290-3,430 sites per cell). The data in Table 1 show that all three types of ¹²⁵I-IL-1 detected a single class of high affinity sites on the cells. Similar experiments with primary human dermal fibroblasts and murine 3T3 fibroblasts also showed that both IL-1 α and IL-1 β detected a single class of cell surface binding sites; the numbers of receptors detected by the two ligands were the same within experimental error (Table 1). In ARH77 cells, nIL-1 β and rIL-1 β bound to similar numbers of high affinity binding sites with identical affinities (Table 1). We could not obtain interpretable

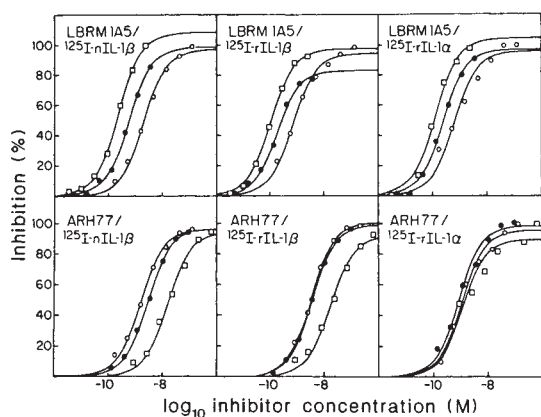


Fig. 2 Cross competition between rIL-1 α , rIL-1 β and nIL-1 β on LBRM-33-1A5 and ARH77 cells. Unlabelled competitor: ●, nIL-1 β ; ○, rIL-1 β ; □, rIL-1 α . LBRM-33-1A5 cells (3.1×10^7 per ml) or ARH77 cells (2.1×10^7 per ml) were incubated with a fixed concentration of each type of 125 I-IL-1 and various concentrations of unlabelled IL-1 (as indicated in the figure) in 150 μ l of RPMI 1640 with 1% BSA, 0.1% sodium azide and 20 mM HEPES pH 7.4 for 2 h at 8°C on a rocker platform in round bottomed 96 well microtitre plates. Concentrations of 125 I-IL-1 used were: with LBRM-33-1A5 cells—nIL-1 β , 3.9×10^{-10} M; rIL-1 β , 1.9×10^{-9} M and rIL-1 α , 9.2×10^{-10} M; with ARH77 cells—nIL-1 β , 9×10^{-10} M; rIL-1 β , 1.9×10^{-9} M and rIL-1 α , 4.6×10^{-9} M. At the end of the assay, cells and bound radiolabel were separated from unbound label by centrifuging 60 μ l aliquots of the reaction mixtures through 200 μ l of phthalate oils in 400 μ l polyethylene centrifuge tubes and excising the tips with a razor blade¹⁷. Maximal binding was measured with the 125 I ligands alone; background binding was measured with 125 I-IL-1 and 6×10^{-7} M unlabelled homologous IL-1. All data have been corrected for nonspecific binding by subtraction. The cells were greater than 90% viable by trypan blue exclusion at the end of the assay. The continuous curves passing through the data were calculated from the best fit parameter values in Table 1 using a simple competitive inhibition model¹⁹.

Table 1 Binding parameters for the interaction of rIL-1 α , rIL-1 β and nIL-1 β with murine and human cells

Cell	Ligand	K_A (M^{-1}) $\times 10^{-9}$	K_i (M^{-1}) $\times 10^{-9}$	Receptors per cell ($\times 10^3$)
LBRM-33-1A5	nIL-1 β	1.4 ± 0.1	4.0 ± 1.8	2.8 ± 0.1
LBRM-33-1A5	rIL-1 β	0.3 ± 0.02	1.8 ± 0.3	3.2 ± 0.2
LBRM-33-1A5	rIL-1 α	4.6 ± 0.7	11.3 ± 2.1	3.43 ± 0.02
ARH77	nIL-1 β	0.44 ± 0.06	0.7 ± 0.4	11.3 ± 0.3
ARH77	rIL-1 β	0.5 ± 0.1	1.5 ± 0.6	11 ± 1
ARH77	rIL-1 α	ND	0.4 ± 0.1	ND
Human dermal fibroblasts	nIL-1 β	1.8 ± 0.1	3.4 ± 0.6	2.4 ± 0.1
Human dermal fibroblasts	rIL-1 α	1.6 ± 0.1	4.1 ± 0.4	2.8 ± 0.1
BALB-c/3T3 fibroblasts	nIL-1 β	2.1 ± 0.6	2.3 ± 1.1	4.8 ± 0.5
BALB-c/3T3 fibroblasts	rIL-1 α	3.0 ± 0.5	2.3 ± 1.1	5.5 ± 2.0

LBRM-33-1A5 murine T lymphoma cells² and ARH77 human B lymphoma cells were grown in 175 cm² tissue culture flasks in RPMI 1640 containing 10% fetal calf serum (Hyclone). Human dermal fibroblasts were obtained from the American Type Tissue Culture Collection (CRL 1229, passage 5) and propagated in monolayer cultures in DMEM 10% fetal calf serum; cells were tested for IL-1 binding at passage 10–12. BALB-c/3T3 cells were obtained from the American Type Tissue Culture Collection and propagated in DMEM with 10% calf serum (Hyclone). Fibroblasts were harvested for binding assays by treatment with PBS containing EDTA (0.5 mM) to avoid proteolysis of receptor. Equilibrium binding data were analysed using a simple bimolecular binding model¹². The values for nonspecific binding were determined by fitting a straight line passing through the origin to data obtained from binding of 125 I-ligand to the cells in the presence of 100-fold molar excess of unlabelled IL-1 of the same type. These values were subtracted from the data before analysis for determination of K_A and receptors per cell values. Inhibition data were analysed with an equation describing competition between two ligands for one class of sites¹⁹. All data analysis was done using RS/1, a commercially available scientific programming package running on a VAX 11/750 under the VMS operating system (BBN Software Products). Error values are standard errors of the mean generated from the nonlinear least-squares fitting program.

revealed that IL-1 α and IL-1 β bind to identical receptor populations (data not shown). Interestingly, the pattern of relative reactivity differed from one cell line to another; IL-1 α bound more avidly than IL-1 β to murine T cells but less avidly to human B cells; both forms bound equally well to fibroblasts, irrespective of the species of origin. The data suggest that the receptors on different cell lineages, although similar, may not be identical. Overall, however, our data are consistent with the view that there are only limited differences in receptor binding activities of IL-1 α and IL-1 β , and almost no difference in receptor binding between nIL-1 β and rIL-1 β .

The observation that the two forms of IL-1 bind to the same receptor on a variety of cells explains their shared biological activities. It is striking that this should occur despite the limited sequence homology between the molecules. This may indicate that the tertiary structures are less divergent than the primary sequences, particularly in the active site region. This feature of the IL-1 system, that IL-1 α and IL-1 β bind to the same receptor protein, is also found for the TNF system²⁰ and the A chain and B chain dimers of PDGF²¹. These three hormone-receptor systems appear to have at least partially overlapping roles in controlling connective tissue metabolism and inflammatory responses. It is therefore tempting to speculate that the existence of two hormones with a common receptor in all three cases is a consequence of some as yet undiscovered feature of the regulation of responses to disease and injury.

Received 16 May; accepted 10 September 1986.

- Gery, I. & Waksman, B. H. *J. exp. Med.* **136**, 143–155 (1972).
- Gillis, S. & Mizel, S. B. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1133–1137 (1981).
- Dewhirst, F. E., Stashenko, P. P., Mole, J. E. & Tsurumachi, T. *J. Immun.* **135**, 2562–2568 (1985).
- Schmidt, J. A., Mizel, S. B., Cohen, D. & Green, I. J. *Immunology* **118**, 2177–2182 (1982).
- Dinarelli, C. A. *Rev. infect. Dis.* **6**, 51–95 (1984).
- Sakatvala, J., Sarsfield, S. J. & Townsend, Y. J. *J. exp. Med.* **162**, 1208–1227 (1985).
- Dinarelli, C. A., Goldin, N. P. & Wolff, S. M. *J. exp. Med.* **139**, 1369–1381 (1974).
- March, C. J. *et al. Nature* **315**, 641–646 (1985).

saturation binding data from these cells for 125 I-labelled rIL-1 α because the affinity of this ligand for the cells is low (see Fig. 2).

To test directly whether IL-1 α and IL-1 β bind to the same receptor we conducted inhibition experiments (Fig. 2) in which the binding of each radiolabelled IL-1 preparation to cells was inhibited by unlabelled IL-1 from the same stock solution and from the other two unlabelled IL-1 preparations. For each 125 I-labelled IL-1 preparation on either LBRM-33-1A5 or ARH77 cells, complete inhibition of radiolabelled ligand binding was achieved with all three unlabelled IL-1 preparations. Thus no receptors on either cell type bound IL-1 α or IL-1 β exclusively. The data in Fig. 2 were analysed quantitatively by a nonlinear least-squares fitting of an equation describing competition between labelled and unlabelled hormone for a single class of sites^{12,19}. The parameter values from this analysis are given in Table 1.

The dose-response curves in Fig. 2 provide further support for the view that both cell types have a single class of IL-1 receptors that bind all three IL-1 preparations. If there existed a subset of receptors that bound one type of IL-1 preferentially, competition between the labelled and unlabelled forms of this type of IL-1 would show a shift to lower concentration of inhibitor than when the unlabelled form competed against labelled heterologous IL-1. This is not observed. Finally, for all 18 data sets in Fig. 3 the curves pass close to most of the data points. A model incorporating only one type of cell surface receptor site thus provides a satisfactory description of the data in all cases.

The results of similar experiments with human and murine fibroblasts are summarized in Table 1. Both for 3T3 cells and for human dermal fibroblasts, cross competition experiments

9. Auron, P. E. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 7907-7911 (1984).
10. Lomedico, P. T. *et al. Nature* **312**, 458-462 (1984).
11. Furutani, Y. *et al. Nucleic Acids Res.* **13**, 5869-5882 (1985).
12. Dower, S. K. *et al. J. exp. Med.* **162**, 501-515 (1985).
13. Dower, S. K., Call, S., Gillis, S. & Urdal, D. L. *Proc. natn. Acad. Sci. U.S.A.* **83**, 1060-1065 (1986).
14. Giri, J. G., Lomedico, P. T. & Mizel, S. B. *J. Immun.* **134**, 343-349 (1985).
15. Kronheim, S. R. *et al. J. exp. Med.* **161**, 490-502 (1985).
16. Schmidt, J. A. *J. exp. Med.* **160**, 772-787 (1984).
17. Segal, D. M. & Hurwitz, E. *J. Immun.* **118**, 1338-1347 (1977).
18. Conlon, P. J. *J. Immun.* **131**, 1280-1282 (1983).
19. Cuatrecasas, P. & Hollenberg, M. D. *Adv. Protein Chem.* **30**, 251-451 (1975).
20. Aggarwal, B. B., Eessalau, T. E. & Haas, P. E. *Nature* **318**, 665-667 (1985).
21. Betsholtz, C. *et al. Nature* **320**, 695-699 (1986).

Presence of Ti (WT31) negative T lymphocytes in normal blood and thymus

Lewis L. Lanier* & Arthur Weiss†

* Becton Dickinson Monoclonal Center, Inc., Mountain View, California 94043, USA

† Department of Medicine, Howard Hughes Medical Institute, University of California, San Francisco, California 94143, USA

The antigen receptor expressed on most T lymphocytes is a disulphide-linked heterodimer (Ti) that is composed of α -chain and β -chain subunits¹⁻⁶. On the surface of human T lymphocytes, Ti is non-covalently associated with three invariant proteins, designated CD3- γ , δ , and ϵ ⁷. It has been suggested that Ti is obligatory for CD3 expression^{8,9}. But a T leukaemia cell line¹⁰, IL-2 (interleukin 2) dependent T-cell clones established from fetal blood¹¹ and IL-2 dependent cell lines established from immunodeficiency patients with bare lymphocyte syndrome and ectodermal dysplasia syndrome¹² have recently been shown to express CD3, but not Ti¹³ (detected due to monoclonal antibody WT31). These lymphocytes may express the product of the T-cell antigen receptor γ (TCR- γ) gene, rather than the α/β heterodimer, in association with CD3^{10,12}. Preliminary studies suggested that T cells expressing CD3 but lacking Ti are present in low frequency in normal lymphoid tissues^{10,12}. Here we show that in normal blood and thymus CD3⁺, WT31⁻ T cells express neither CD4 nor CD8. The low frequency (<0.2-0.9% of total thymocytes) of CD3⁺, WT31⁻ cells in the thymus suggests that this population does not represent a major stage of thymic development and may be a distinct lineage of T cells.

Peripheral blood lymphocytes (PBL) were stained by direct two-colour immunofluorescence with fluorescein isothiocyanate (FITC) conjugated anti-Ti (WT31) and biotin conjugated anti-CD3 (Leu 4), followed by phycoerythrin (PE) conjugated streptavidin and analysed by flow cytometry. Although most CD3⁺ T cells co-expressed Ti (WT31), a small, distinct subset of CD3⁺ T cells did not. Based on analysis of >30 normal adult blood donors, ~3% of T lymphocytes were CD3⁺, WT31⁻ (range <0.5%-10%), whereas the remainder were CD3⁺, WT31⁺. CD3 is expressed on all peripheral blood T cells, but these cells can be divided into subpopulations on the basis of expression of other surface antigens. For example, the CD4 and CD8 antigens are mutually exclusive markers for the two major subpopulations of mature T cells. This phenotypic heterogeneity has some functional correlation; CD4-bearing cells recognize antigen in association with class II major histocompatibility (MHC) antigens and generally function as helper cells, whereas the CD8 bearing cells recognize antigen in association with class I MHC antigens and generally function as cytotoxic T cells. Thus it was of interest to determine whether CD3⁺, WT31⁻ cells could be detected in one of these major subpopulations, as this would have implications for cells bearing this unusual CD3 receptor complex. Three-colour immunofluorescence was performed to

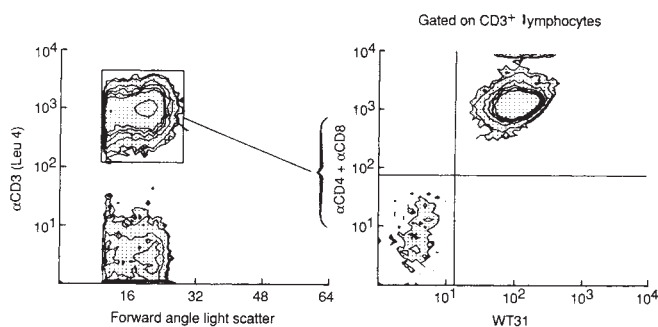


Fig. 1 Presence of CD3⁺, WT31⁻ T lymphocytes in peripheral blood. Mononuclear cells were isolated from peripheral blood using Ficoll/Hypaque. Adherent cells were removed using plastic tissue culture flasks. PBL were stained with FITC-conjugated WT31, biotin-conjugated anti-CD3 (Leu 4) and a mixture of PE-conjugated anti-CD4 (Leu 3) and anti-CD8 (Leu 2). APC-conjugated streptavidin was used as the second reagent. Fluorochrome-conjugated, isotype-matched monoclonal antibodies that do not react with human cells were used to control for nonspecific immunofluorescence. Samples were analysed by multi-parameter flow cytometry using a dual laser FACS 440 system with an argon ion laser (488 nm; 100 mW) and a helium neon laser (633 nm; 33 mW) (Becton Dickinson Immunocytometry Systems). Immunofluorescence, flow cytometry and data analysis were performed as described^{14,15}. Based on the immunoglobulin control sample, contour plots were divided into quadrants to identify unstained cells (lower left); cells stained with both fluorochromes (upper right), and cells stained only with a single fluorochrome (upper left and lower right). In the control sample >99% of lymphocytes were unstained (lower left quadrant, not shown). Left, forward angle light scatter (linear scale, proportional to cell size for lymphoid cells) versus APC anti-CD3 fluorescence. Insert, An 'electronic' gate on CD3⁺ T lymphocytes. The correlation of the WT31, CD4 and CD8 antigens on CD3⁺ T lymphocytes is displayed (right panel). In this individual 8% of total lymphocytes were CD3⁺WT31⁻ and all lacked CD4 and/or CD8 expression.

determine whether CD3⁺, WT31⁻ T lymphocytes in blood co-express the CD4 or CD8 antigens. Surprisingly, they express neither CD4 nor CD8 (Fig. 1). By setting an electronic gate on CD3⁺ lymphocytes (stained with allophycocyanin (APC)-conjugated streptavidin/biotin-conjugated anti-CD3), it was possible to examine directly CD3⁺ T cells co-stained with FITC-conjugated anti-Ti (WT31) and a mixture of PE-conjugated anti-CD4 plus PE-conjugated anti-CD8 (Fig. 1, right panel). Note that: (1) all CD3⁺, WT31⁻ T lymphocytes lacked CD4 and CD8, and (2) all detectable CD3⁺, WT31⁻ T cells expressed either CD4 or CD8. Identical results were obtained from several independent donors.

Further experiments were conducted to determine whether this unique subset of CD3⁺, WT31⁻ T cells expressed other T-cell-associated differentiation antigens. As shown in Fig. 2, >95% of CD3⁺, WT31⁻ T lymphocytes co-expressed CD2, CD5 and CD7, similar to expression of these antigens on CD3⁺, WT31⁺ T lymphocytes. CD1 (a class I MHC homologous antigen limited to thymocytes), CD19 (a pan B cell associated antigen), CD14 (a monocyte associated antigen), transferrin receptors and IL-2 receptors were not expressed on CD3⁺, WT31⁻ T cells (not shown).

We have shown previously¹⁰ that CD3⁺, WT31⁻ cells comprise <1% of total thymocytes. As blood CD3⁺, WT31⁻ T cells are CD4⁻8⁻, we examined CD4⁻8⁻ thymocytes (6% of total thymocytes) for expression of CD3 or WT31. As shown in Fig. 3 (top), essentially all thymocytes expressing CD4 and/or CD8 showed linked expression of CD3 and Ti; cells were either CD3⁺, Ti⁺ or CD3⁻, Ti⁻. The 'diagonal' in the bivariate contour plot indicates a direct quantitative relationship between CD3 and Ti. Further analysis by three-colour immunofluorescence

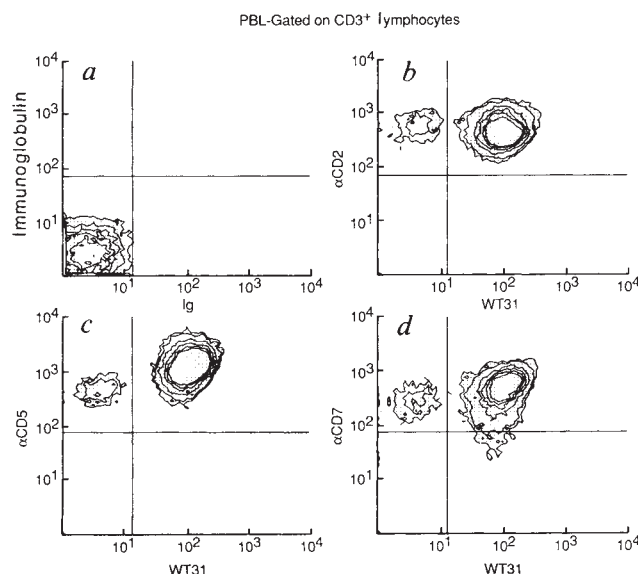


Fig. 2 Expression of T-cell differentiation antigens on $CD3^+$, $WT31^-$ T lymphocytes. PBL (from the same individual as in Fig. 1) were stained with: *a*, biotin-conjugated anti-CD3 (Leu 4), FITC-conjugated immunoglobulin control, and PE-conjugated immunoglobulin control; *b*, biotin anti-CD3, FITC-conjugated WT31, and PE-conjugated anti-CD2 (Leu 5); *c*, biotin-conjugated anti-CD3, FITC-conjugated WT31, and PE-conjugated anti-CD5 (Leu 1); *d*, biotin-conjugated anti-CD3, FITC-conjugated WT31, and PE-conjugated anti-CD7 (Leu 9). APC-conjugated streptavidin was used as the second step reagent. An 'electronic' gate was placed on $CD3^+$ T lymphocytes (shown in Fig. 1, left) and the correlation of WT31 with CD2 (B), CD5 (C) or CD7 (D) was directly determined. In control samples (using fluorochrome conjugated isotype-matched immunoglobulin) >99% of lymphocytes were unstained.

(staining with FITC-conjugated anti-CD8, PE-conjugated anti-CD4 and APC-conjugated anti-CD3)¹⁴ indicated that essentially all 'single positive' $CD4^+8^-$ and $CD4^+8^+$ and about half the immature $CD4^+8^+$ thymocytes expressed CD3 (not shown). By contrast, a very minor subset of $CD4^+8^-$ thymocytes expressed CD3 but lacked Ti (WT31) (0.8% of total thymocytes; Fig. 3). In fact, all detectable $CD3^+$ lymphocytes in this population were Ti (WT31) negative. Furthermore, like peripheral blood $CD3^+$, $WT31^-$ lymphocytes, $CD3^+4^-8^-$ thymocytes expressed CD2 and CD5, but not transferrin receptors or IL-2 receptors (Fig. 4). In contrast to peripheral $CD3^+$, $WT31^-$ lymphocytes, CD1 was

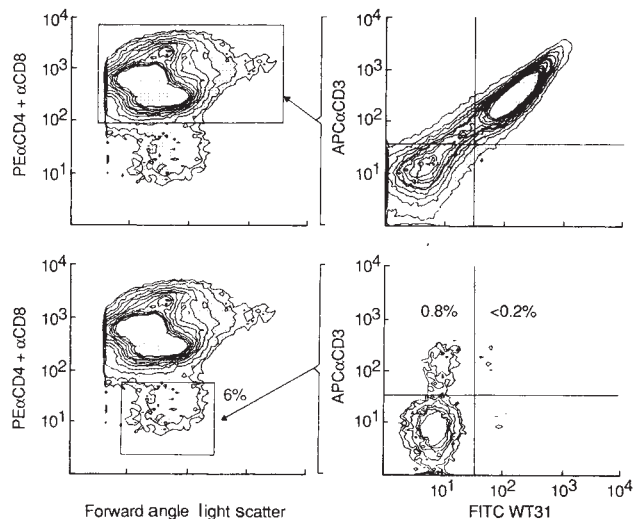


Fig. 3 Expression of CD3 and WT31 on thymocytes. Normal thymus were obtained from paediatric cardiac patients. A single cell suspension of thymocytes was prepared, using the entire thymic lobule to ensure that cortical and medullary thymocytes were accurately represented, and stained with FITC-conjugated WT31, biotin-conjugated anti-CD3 (Leu 4) and a mixture of PE-conjugated anti-CD4 (Leu 3) with anti-CD8 (Leu 2), followed by APC-conjugated streptavidin. An electronic gate was set on thymocytes expressing CD4 and/or CD8 (top) or lacking both CD4 and CD8 (bottom) and the correlation between WT31 and CD3 was determined. In this individual 6% of thymocytes were $CD4^+8^-$ and 0.8% were $CD3^+4^-8^-$, $WT31^-$. In samples stained with fluorochrome-conjugated, isotype-matched control Ig <0.2% of thymocytes were nonspecifically stained (not shown).

expressed on a fraction (~40%) of these $CD3^+$, $WT31^-$ thymocytes (Fig. 4d), although the amount of CD1 antigen was approximately 10% the amount present on most thymocytes (not shown). Thymocytes from five donors (age 2–6 years) were examined with similar results (mean ~0.5% $CD3^+$, $WT31^-$ thymocytes; range <0.2–0.9%).

We have thus shown that $CD3^+$ T lymphocytes lacking Ti are present in normal blood and at a lower frequency in the thymus. $CD3^+$, $WT31^-$ T cells and thymocytes are only found in the small $CD4^+8^-$ subset. Peripheral blood $CD3^+4^-8^-$ T lymphocytes are morphologically and phenotypically mature, small resting lymphocytes¹⁵. IL-2 dependent cell lines of $CD3^+4^-8^-$ thymocytes¹⁶ and normal peripheral blood T lymphocytes¹⁶ established in culture stably express this phenotype, remain Ti

Gated on $CD4^+8^-$ thymocytes

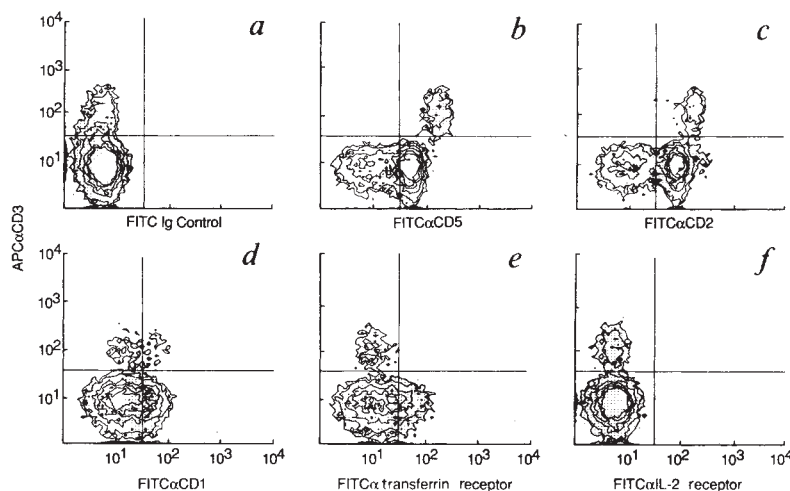


Fig. 4 Expression of T-cell differentiation antigens on $CD3^+4^-8^-$ thymocytes. Thymocytes (from the same individual as in Fig. 3) were stained with biotin-conjugated anti-CD3 (Leu 4), a mixture of PE-conjugated anti-CD4 (Leu 3) and anti-CD8 (Leu 2), and; *a*, FITC-conjugated Ig control; *b*, anti-CD5 (Leu 1); *c*, anti-CD2 (Leu 5); *d*, anti-CD1 (Leu 6); *e*, anti-transferrin receptor; *f*, anti-IL-2 receptor. APC-conjugated streptavidin was used as the second step reagent. As shown in Fig. 3 (bottom panel), an electronic gate was set on $CD4^+8^-$ thymocytes and the correlation between expression of CD3 and of the other differentiation antigens was determined.

(WT31) negative, and demonstrate cytotoxic function (refs 16, 17 and L.L.L., unpublished). Some cell lines with this phenotype^{12,16} do not express mature Ti- α or Ti- β messenger RNA but express a novel heterodimer (probably composed of TCR- γ and another unidentified glycoprotein) in association with CD3. The CD3⁺ WT31⁻ cells we have observed *in vivo* are CD4⁺ 8⁻, but some IL-2 dependent lines with this phenotype express CD8^{11,12}. We do not know whether CD3⁺ 8⁺ WT31⁻ cells are present *in vivo* in numbers we cannot detect (<0.2%) or whether CD8 arises on these cells as a consequence of culture.

These observations raise questions regarding the lineage and differentiation pathway of these unique CD4⁺ 8⁻ T lymphocytes. As rearrangement and transcription of TCR- γ genes apparently precedes Ti- β and Ti- α ¹⁷⁻²¹, it has been proposed that the antigen receptor on immature T cells may be composed of TCR- γ gene products before the Ti- α / β heterodimer appears. However because CD3⁺, WT31⁻ thymocytes are a minor subpopulation of the putative immature CD4⁺ 8⁻ thymocytes, it is unlikely that CD3⁺, WT31⁻ T cells represent a major stage in thymic development. If selection of antigen/MHC recognition is from a CD3⁺, WT31⁻ thymic subset, there should be more of these cells. Because most thymocytes (including the abundant immature CD4⁺ 8⁺ subset) express Ti (WT31), selection may occur mainly via the CD3/Ti receptor complex. Alternatively, CD3⁺ Ti⁻ lymphocytes may be a distinct T-cell lineage and never express Ti. The fact that CD3⁺, WT31⁻ cells in blood are identical in phenotype to the thymocyte subpopulation (except for CD1, found only on thymocytes) supports this possibility. Finally, CD3⁺, WT31⁻ peripheral blood cells could be precursors of the CD3⁺, WT31⁻ thymocytes which might then be important in thymocyte development. If so, a cell population equivalent to the CD3⁺, Ti⁻ T cells should be present in athymic mice, and might account for the T-cell-mediated functions observed in these animals²²⁻²⁵. MacDonald *et al.*²⁶ have recently described a subset of Thy1⁺, L3T4⁻, Ly2⁻ cells in athymic mice and Mak²⁷ has recently reported that nylon wool non-adherent lymphocytes from athymic mice transcribe TCR- γ , but not Ti- α or competent Ti- β mRNA. Further studies are necessary to test this hypothesis and establish the relationship between CD3⁺, Ti⁺ and CD3⁺, Ti⁻ T lymphocytes.

We thank Drs Willem Tax and Patrick Kung for providing the WT31 antibody and Dr Jim Allison for discussion and review of the manuscript.

Received 1 August; accepted 8 September 1986.

- Allison, J. P., McIntyre, B. W. & Bloch, D. *J. Immun.* **129**, 2293-2300 (1982).
- Haskins, K. *et al. J. exp. Med.* **157**, 1149-1169 (1983).
- Meuer, S. C. *et al. J. exp. Med.* **157**, 705-719 (1983).
- Acuto, O. *et al. Cell* **34**, 717-725 (1983).
- Kapler, J. *et al. Cell* **35**, 295-302 (1983).
- McIntyre, B. & Allison, J. P. *Cell* **38**, 659-665 (1984).
- Borst, J., Prendiville, M. A. & Terhorst, C. *J. Immun.* **128**, 1560-1565 (1982).
- Weiss, A. & Stobo, J. D. *J. exp. Med.* **160**, 1284-1299 (1984).
- Ohashi, P. S. *et al. Nature* **316**, 606-609 (1985).
- Weiss, A., Newton, M. & Crommie, D. *Proc. natn. Acad. Sci. U.S.A.* **83**, 6998-7002 (1986).
- Nowill, A. *et al. J. exp. Med.* **163**, 1601-1606 (1986).
- Brenner, M. B. *et al. Nature* **322**, 145-149 (1986).
- Oettingen, H. C., Kappler, J., Tax, W. J. H. & Terhorst, C. *J. biol. Chem.* **259**, 12039-12048 (1984).
- Lanier, L. L., Allison, J. P. & Phillips, J. H. *J. Immun.* **137**, 2501-2507 (1986).
- Lanier, L. L., Ruitenberg, J. J. & Phillips, J. H. *J. exp. Med.* **164**, 339-344 (1986).
- Bank, I. *et al. Nature* **322**, 179-181 (1986).
- Snodgrass, H. R., Dembic, Z., Steinmetz, M. & von Boehmer, H. *Nature* **315**, 232-233 (1985).
- Raulet, D. H., Garman, R. D., Saito, H. & Tonegawa, S. *Nature* **314**, 103-107 (1985).
- Garman, R. D., Doherty, P. J. & Raulet, D. H. *Cell* **45**, 733-742 (1986).
- Haars, R. *et al. J. exp. Med.* **164**, 1-24 (1986).
- Sangster, R. N., Minowada, J., Sucin-Foca, N., Minden, M. & Mak, T. W. *J. exp. Med.* **163**, 1491-1508 (1986).
- Hunig, T. *Immun. Today* **4**, 84-87 (1983).
- Gillis, S., Union, N. A., Baker, P. E. & Smith, K. A. *J. exp. Med.* **149**, 1460-1476 (1979).
- Hunig, T. & Bevan, M. J. *J. exp. Med.* **152**, 688-702 (1980).
- MacDonald, H. R. & Lees, R. K. *J. Immun.* **132**, 605-610 (1984).
- MacDonald, H. R., Blanc, C., Lees, R. K. & Sordat, B. *J. Immun.* **136**, 4337-4339 (1986).
- Mak, T., Yoshikawa, Y. & Reis, M. D. *Nature* (in the press).
- Parks, D. R., Lanier, L. L. & Herzenberg, L. A. in *Handbook of Experimental Immunology*, 4th edn (eds Weir, D. M. *et al.*) 29.1-29.2 (Blackwell, Oxford, 1986).

Structure of pre-pro-von Willebrand factor and its expression in heterologous cells

David T. Bonthron*†, Robert I. Handin‡, Randal J. Kaufman§, Louise C. Wasley§, Elizabeth C. Orr§, Lisa M. Mitsock§, Bruce Ewenstein‡, Joseph Loscalzo‡, David Ginsburg*‡|| & Stuart H. Orkin*†

* Division of Hematology/Oncology, Children's Hospital, 300 Longwood Avenue, Boston, Massachusetts 02115, USA and the Dana-Farber Cancer Institute, 44 Binney Street, Boston, Massachusetts 02115, USA

† Howard Hughes Medical Institute and the Departments of Medicine and Pediatrics, Harvard Medical School, 25 Shattuck Street, Boston, Massachusetts 02115, USA

‡ Division of Hematology, Brigham and Women's Hospital, 75 Francis Street, Boston, Massachusetts 02115, USA

§ Genetics Institute, 87 Cambridge Park Drive, Cambridge, Massachusetts, USA

Von Willebrand factor (vWF), a multifunctional haemostatic glycoprotein derived from endothelial cells and megakaryocytes, mediates platelet adhesion to injured subendothelium and binds coagulation factor VIII in the circulation. Native vWF is a disulphide-bonded homopolymer; the monomeric subunits, of apparent relative molecular mass (M_r) 220,000 (220K) are derived from an intracellular precursor estimated at 260-275K (refs 1-5). Multimer assembly is preceded by the formation of dimers, linked near their C-termini, which then assemble into filamentous polymers. The importance of the removal of the large vWF pro-polypeptide during multimer assembly, and whether this or other stages of the complex post-translational processing require components specific to endothelial cells or megakaryocytes, is unknown. Here we report an analysis of the complete sequence of pre-pro-vWF and expression of the molecule in heterologous cells. The vWF precursor is composed of several repeated subdomains. When expressed in COS and CHO cells, it is cleaved and assembled into biologically active high relative molecular mass disulphide bonded multimers. This suggests that the information for assembly of this complex molecule resides largely within its primary structure.

The nucleotide sequence of pre-pro-vWF complementary DNA (which is printed in its entirety elsewhere⁴⁸) has a major open reading frame encoding 2,813 amino acids. The initiator methionine, followed by a classical signal peptide, can be identified unambiguously because it is the only Met between an in-frame upstream stop codon and the Ala (residue 23) determined by amino-acid sequencing to be the N-terminus of the 100K pro-polypeptide⁶. The signal peptidase cleavage site (defined by this N-terminal sequence) conforms well to the consensus⁷. The predicted M_r of pro-vWF is 307K without the extra 15% contributed by glycosylation⁸; this is considerably larger than the size of the intracellular precursor estimated from SDS-PAGE⁵.

The amino-acid sequence of pre-pro-vWF shows extensive internal repetition (Fig. 1). A central region of ~600 amino-acids contains three domain A repeats, which together with the short B and C repeats, were identified in partial-length clones described by others^{9,10}. A striking feature of the A region is its lack of cysteine residues; in contrast, the flanking regions (amino acids 34-1,242 and 1,948 to the C-terminus) are rich in cysteine, the most abundant amino acid in pre-pro-vWF, comprising 8.3% of the residues (10.4% in the regions outside domain A). These cysteine-rich regions include a further series of four repeats, which we term D1-D4, beginning just after the signal peptide (Fig. 1). The second of these 350 amino-acid repeats is truncated,

|| Present address: Howard Hughes Medical Institute, 1150 West Campus Drive, Ann Arbor, Michigan 48109, USA.

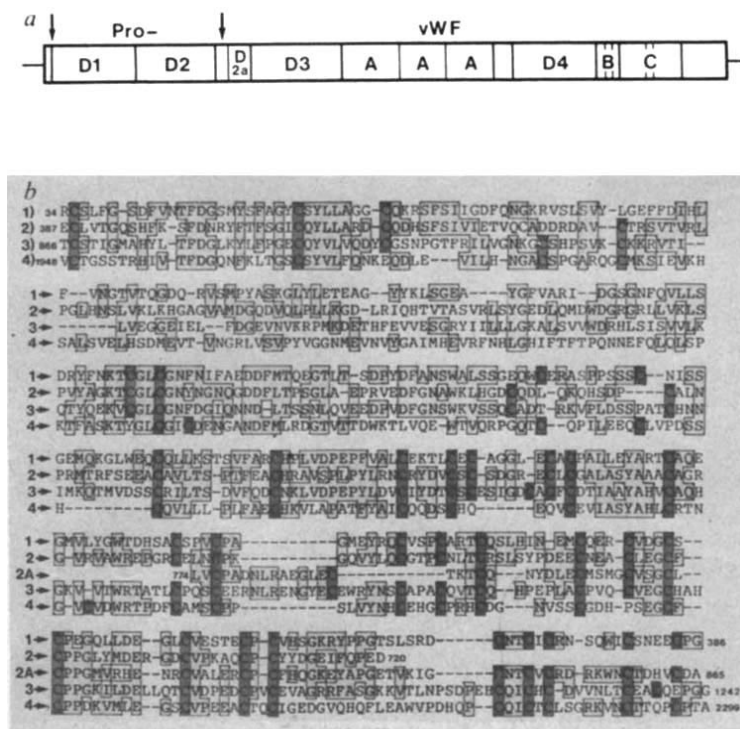


Fig. 1 Repeated elements in the primary structure of vWF. *a*, Schematic of vWF cDNA. D1–D4, cysteine-rich repeat blocks. The A, B and C repeat domains of Sadler *et al.*^{9,10} are also labelled. Arrows, sites at which the signal peptide and the large pro-polypeptide are cleaved from the primary translation product. *b*, Alignment of the four cysteine-rich D repeat blocks. Matching residues are boxed and lightly shaded, cysteine residues are darkly shaded. Small numbers at the sides refer to the adjacent amino-acid residues, relative to the initiator methionine, assigned position 1. Amino acids 721–773, (between repeats 2 and 2A), and 1,243–1,947 (including the A domain repeats), have been omitted; the sequence is otherwise continuous as shown. Gaps introduced to maximize homology are shown as dashes. D4 appears to be the least conserved of the repeat blocks; in this alignment, the percentage of matching amino acids for each pairwise comparison is: D1–D2, 34.8; D1–D3, 30.4; D2–D3, 33.6; D1–D4, 24.1; D2–D4, 25.0; D3–D4, 24.1. The sequence TCGLC, which forms the centre of a highly conserved block in the D repeat regions, also occurs in the B domain, indicating a possible relationship between these two repeated elements.

Methods. The amino-acid sequence shown is deduced from the sequences of the same cDNA clones used in construction of pvWF2 (Fig. 2 legend). Both strands of the sequence were determined by the Sanger dideoxy method^{34,35}. M13 subclones for sequencing were constructed by exonuclease digestion methods^{36,37}. The region of residue 1,472 (see text) was sequenced in three further independent clones (pvWG8b, pvWH22, pvWH23) using an internal oligonucleotide primer³⁸. Repeat blocks were defined manually following identification of short repeated motifs by computer search. The consensus alignment was also arrived at by hand, with reference to computer-generated pairwise alignments.

and its C-terminal one-third duplicated (Fig. 1, D2a). Interestingly, the site of pro-polypeptide cleavage (Fig. 1, arrow) falls within this region of discontinuity of the D repeats.

Based on partial cDNA sequences, Sadler and colleagues¹⁰ proposed the existence of a fifth repeated element which they term domain E. Our analysis shows that this element is, in fact, the C-terminal portion of the D repeat block, and these sequences are included in the alignment in Fig. 1.

Within the D repeats there is striking conservation of the positions of the abundant cysteine residues (Fig. 1*b*). Because no free sulphhydryl groups can be detected in native vWF^{11,12} all these cysteines must participate in inter- and intra-chain disulphide bonding, an important determinant of the tertiary and quaternary structure of the protein. Electron microscopy of vWF shows a central filament separating globular regions at each end of the monomer^{13–15}. These two structural features may correspond to the cysteine-poor A domains and the highly disulphide-bonded cysteine-rich domains, respectively. The intron–exon arrangement underlying the repeated structure of vWF has not yet been defined, but the common evolutionary origin of mature vWF and the pro-polypeptide is evident from our analysis. As the cleavage of pro-vWF occurs at a position where the cysteine-rich repeat sequence is interrupted by a partial duplication, it could be that after the duplications of the D repeats a later aberrant recombination introduced the new cleavage site, leading to the separation of the pro-piece from mature vWF. Despite the functional similarities between vWF and fibronectin¹⁶ and fibrinogen¹⁷, we find no similarity in sequence between these proteins other than the consensus

tetrapeptide Arg-Gly-Asp-Ser, which defines the binding site for platelet glycoprotein IIb/IIIa and is common to many adhesive glycoproteins^{18–20}.

Discrepancies exist between our full-length sequence and the partial sequence of Sadler *et al.*^{9,10}. The first 13 base pairs (bp) of the clone λ HvWF-1 (ref. 9) differ completely from our sequence, which in this region was determined on two independent clones, pvWH33 and pvWH5. At amino-acid residue 770 (residue 7 of the mature monomer), the Sadler *et al.* sequence has His rather than the Pro determined from N-terminal amino-acid sequencing⁹, whereas our clones H33 and H5 both have the expected Pro. At residue 1,472, our sequence contains CAC (His) rather than GAC (Asp). Of three further non-identical clones from our endothelial cDNA library, two have GAC and one CAC at this position, suggesting an amino-acid polymorphism. Some other potential polymorphisms have been previously noted¹⁰.

To study vWF biosynthesis in heterologous cells, we assembled a full-length cDNA and inserted it into the expression vector pMT2 (Fig. 2 legend), in which transcription is directed by the adenovirus major late promoter. In assays of transient expression, the conditioned medium of monkey kidney (COS) cells transfected with the pMT2–vWF construct contained 50–100 ng ml^{−1} of vWF antigen. COS cells transfected with pMT2 alone did not produce material reacting in our inhibition enzyme-linked immunosorbent assay (ELISA).

For expression of vWF in Chinese hamster ovary (CHO) cells, we used a second expression vector, pMT2ADA–vWF (Fig. 2) with a protocol of selection for cells over-expressing the enzyme

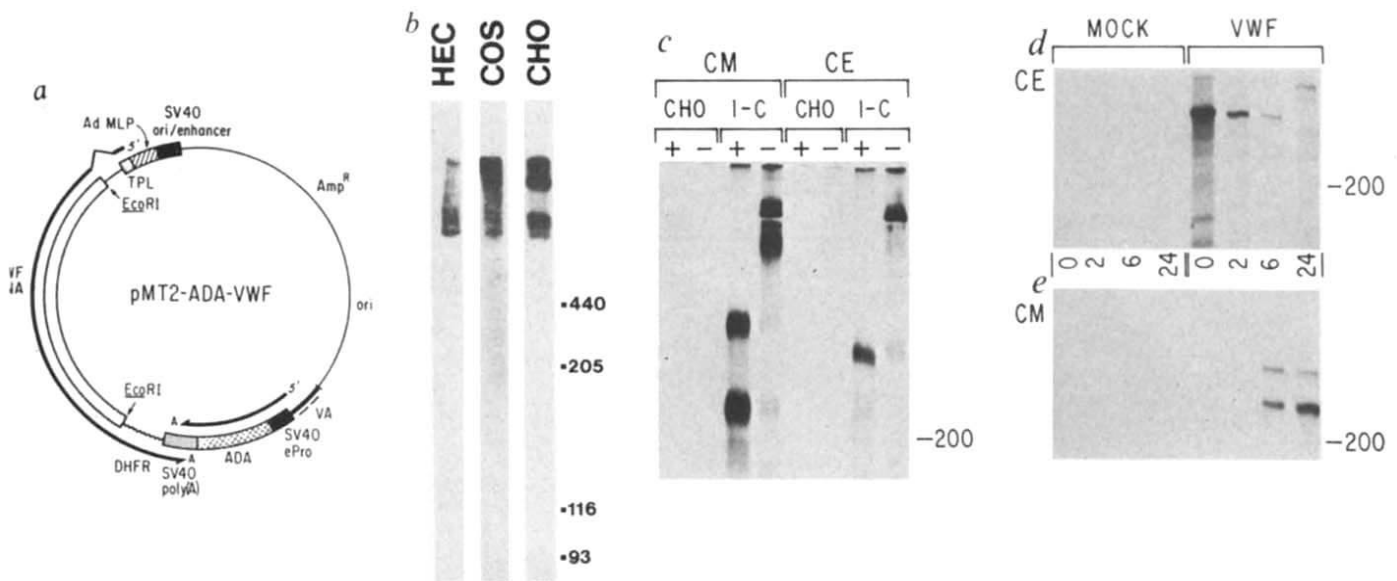


Fig. 2 Expression of vWF cDNA. *a*, Plasmid pMT2ADA-vWF used for expression in CHO cells. It is similar to pMT2-vWF, but contains in addition a human adenosine deaminase (ADA) transcription unit. AdMLP, adenovirus major late promoter; TPL, tripartite leader. VA genes: ePro; simian virus 40 (SV40) early promoter. *b–e*, Analysis of immunoprecipitated material from cell extracts or conditioned medium of passaged human endothelial cells (HEC), COS cells or CHO cells. *b*, Immunoprecipitates from conditioned medium analysed unreduced on a 4.3% gel (extended run time). *c*, Labelled immunoprecipitates from conditioned medium (CM) or cell extracts (CE) of CHO cells (untransfected) and the transfected, ADA-selected line 1-C, with (+) or without (–) reducing agent. *d*, Pulse-chase of 35 S-labelled material from transfected or mock-transfected COS cell extracts. *e*, Pulse-chased material in conditioned medium from the same experiment.

Methods. Segments of overlapping cDNA clones³² were recombined as follows: pVWH33 from the 5' *EcoRI* linker to the unique *BamHI* site; pVWE6 from the *SacII* site to the 3' *EcoRI* linker; pVWH5 between the above *BamHI* and *SacII* sites. The full-length cDNA in plasmid pUC13³⁹ was designated pVWF2. For transient expression in COS cells, the *EcoRI* fragment containing the entire coding region was transferred to the expression vector pMT2, generating pMT2-vWF. pMT2 is a derivative of the mammalian cell expression vector p91023(B)⁴⁰ in which the ampicillin resistance marker has replaced the tetracycline resistance marker. pMT2-vWF was grown in *Escherichia coli* DH5 (*recA*) to prevent deletion of vWF sequences. The expression vector used for CHO cells was constructed by inserting a *PvuII/HinI* fragment from pSV2 ADA (1a)⁴¹, containing the SV40 early promoter and human ADA coding region, into a unique *XbaI* site in pMT2. This generated a transcription unit that uses the SV40 late polyadenylation signal present in pMT2 for expression of ADA. The vWF cDNA fragment was then inserted at the unique *EcoRI* cloning site to generate pMT2ADA-vWF (*a*). Protoplasts of *E. coli* DH5 carrying pMT2ADA-vWF were fused to CHO cells as described⁴². After 2 days, the cells were subcultured into selective medium containing 1.1 mM adenosine, 1 mM uridine, 50 μ M alanosine and 0.1 μ M 2' deoxycoformycin²⁰. Transformants (~100) were pooled and propagated sequentially in 0.5 and 1.0 μ M 2' deoxycoformycin to select for amplification of the ADA and vWF cDNAs. One such line obtained was designated 1-C. For assay of vWF, these cells were rinsed and incubated for 24 hours in serum-free medium. COS cells were transfected with pMT2-vWF or with pMT2 by the DEAE-dextran technique, with chloroquine treatment^{43–45}. Forty-eight hours after transfection, cells were rinsed and fed serum-free medium for 24 h. Culture medium was then collected. vWF antigen was measured by an inhibition ELISA, using affinity-purified rabbit anti-vWF antiserum, purified vWF antigen and biotinylated second antibody (goat anti-rabbit IgG) with colour development by avidin-peroxidase complex. COS cells were labelled with 35 S-cysteine (0.5 mCi ml⁻¹, NEN) for 15 min, 60 h after transfection. They were then fed with serum-free medium containing cysteine, and incubated at 37 °C for the times indicated. Control CHO cells and the 2' deoxycoformycin-resistant CHO cells were labelled for 3 h with 35 S-methionine (0.5 mCi ml⁻¹), followed by addition of an equal amount of label for a further 3 h. CM and CE were prepared for immunoprecipitation as described⁴⁶. Immunoprecipitation was carried out with a 1:40 dilution of affinity purified rabbit anti-vWF antibody for 8–12 h at room temperature, followed by either Staphylococcal protein A or goat anti-rabbit IgG immobilized on Sepharose beads. After 2 h incubation the beads were washed and taken up in SDS-gel loading buffer with or without 1 mM dithiothreitol, heated to 100 °C and run on 6% (*d–e*) or 4.3% (*b–c*) polyacrylamide gels using the Laemmli buffers⁴⁷, followed by drying and autoradiography.

adenosine deaminase to amplify the plasmid sequences²¹. The cell line obtained (1-C) secretes 4 μ g ml⁻¹ vWF antigen per 10⁶ cells per 24 h as measured by ELISA, a level approaching that of normal plasma (5–10 μ g ml⁻¹).

Functional vWF multimers are normally assembled only in endothelial cells and megakaryocytes via a complex series of processing steps^{3–5}. The vWF pro-polypeptide cleavage site, preceded by the basic amino acids Lys-Arg, is similar to the cleavage sites of many prohormones (see, for example, refs 22–27). Although COS cells can correctly process at least one such site²⁸, the requirements for precursor cleavage have not been defined, and correct removal of the large vWF pro-polypeptide could not be confidently expected. Furthermore, the assembly of vWF monomers into disulphide-bonded dimers and subsequently into linear arrays of high *M_r* multimers might require enzymes or organelles restricted to the specialized cell types that normally secrete vWF.

To investigate the post-translational processing of recombinant vWF, we labelled cells with 35 S Cys or Met and immunoprecipitated vWF from medium and cell extracts. We analysed unreduced immunoprecipitates on 4.3% polyacrylamide gels. A series of high *M_r* bands is seen in medium from passaged human endothelial cells (HEC) and from transfected

COS and CHO cells (Fig. 2*b*), although the lack of resolution in this region of the gel precludes detailed comparison of the multimer structure. On reduction and denaturation, the immunoprecipitated material from the medium (CM) of transfected CHO cells (Fig. 2*c*) or COS cells (not shown) runs as the expected two bands of *M_r* 260K and 220K. Control immunoprecipitates from untransfected CHO cells are negative (Fig. 2*c*, lanes 1, 2). Cell lysates (CE) contain predominantly the precursor species (Fig. 2*c*, lane 7), which appears slightly smaller than the precursor seen in conditioned medium, in accord with previous observations on endothelial cells⁵. In pulse-chase experiments on transfected COS cells (Fig. 2*d, e*, cell extract and conditioned medium, respectively), labelled vWF, predominantly in the processed form, is present in the cell medium from 6 h, while the labelled precursor disappears from the cell extracts over a similar time period.

These studies indicate that cells expressing recombinant vWF not only remove the vWF pro-polypeptide but can assemble high *M_r* disulphide-bonded multimers from the cleaved product. To assess the functional properties of recombinant vWF, we examined its interactions with platelet glycoprotein Ib and with collagen. Binding of native vWF multimers to these two macromolecules results in attachment of platelets to vascular

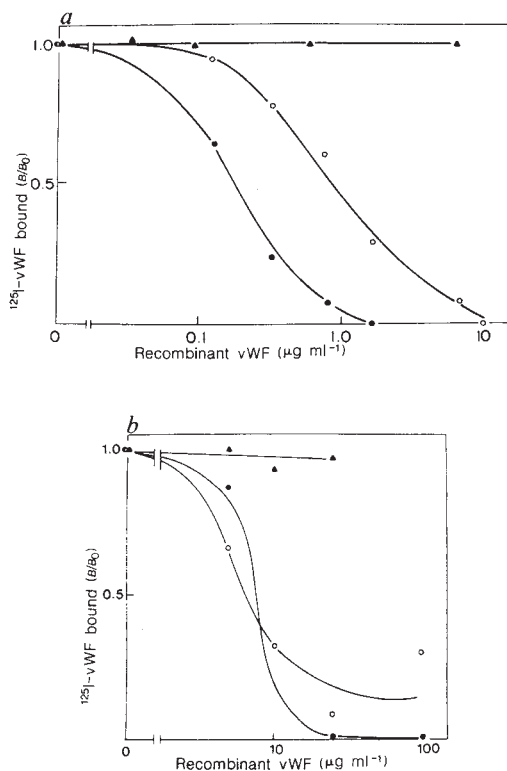


Fig. 3 Biological activity of recombinant vWF. The ability of culture media derived from COS cells transfected with pMT2 or pMT2-vWF to compete with purified native vWF for binding to platelet glycoprotein Ib (GPIb) or to collagen coated monolayers is shown. *a*, Competition for binding to collagen-coated monolayers (closed circles) is compared with competition by purified plasma vWF (open circles). *b*, Competition for binding to GPIb using ristocetin to induce binding to freshly prepared washed platelets (closed circles) or formalin-fixed platelets (open circles). The control supernatants did not compete in either experiment (triangles).

Methods. vWF multimers were purified from normal human plasma and iodinated by the Iodobead technique²⁹. To raise the concentration of recombinant vWF from COS cells sufficiently to carry out the radioligand assays, 50–100 ml conditioned medium was dialysed against 50% Ficoll for 6 to 8 h, and the 10 fold-concentrated samples were reassayed with the inhibition ELISA. For analysis of GPIb binding, platelets were washed free of plasma and used immediately or treated with 1% paraformaldehyde and stored at 4°C for several weeks. Binding of ¹²⁵I-vWF was induced by the addition of 1.5 mg ml⁻¹ final concentration of ristocetin, as previously described²⁹. For the collagen binding assays, a 1 mg ml⁻¹ solution of acid-soluble types I and III collagen at pH 3.0 was added to the wells of an Immulon microtitre plate. Unbound collagen was aspirated after 1 h, and the wells washed with a series of buffered solutions to raise the pH to 7.4. Labelled vWF and competing ligand were added to the wells, and after 1 h at room temperature unbound material was removed by aspiration and several washes with 150 mM NaCl, 10 mM Tris, pH 7.6. The wells were then detached from the plate and bound radioactivity determined. In previous studies binding was found to be specific, saturable and reversible, and to require vWF multimers and native collagen³⁰.

subendothelium. We used concentrated COS cell supernatants as competing ligand against purified, ¹²⁵I-labelled human plasma vWF multimers under conditions in which native vWF binds specifically and saturably (Fig. 3; refs 29, 30). The concentration of recombinant vWF displacing 50% of native vWF from collagen (IC₅₀) is 0.2 μg ml⁻¹, fivefold lower than the IC₅₀ obtained using purified human vWF. For GPIb binding, the IC₅₀ for recombinant vWF (Fig. 3b) and plasma vWF (not shown) are

equivalent (5–7 μg ml⁻¹). Control medium from cells transfected with the vector pMT2 does not compete (Fig. 3a, b, solid triangles). Multimeric recombinant vWF thus exhibits similar activity to native vWF in two relevant bioassays.

The availability of vWF cDNA clones^{31–33} permits new experimental approaches to the complex biosynthesis of this protein. Here, we show that recombinant vWF is processed and secreted in high *M_r* form by two heterologous cell types. The multimer conformation is sufficiently like that of native vWF that the recombinant material competes effectively for binding to platelets and collagen. In fact, recombinant vWF competes more effectively than vWF derived from human plasma for binding to collagen, possibly because some of the highest *M_r* and most active multimers are lost during purification of plasma vWF, whereas recombinant vWF was not fractionated or chromatographed before analysis. This suggests that assembly of vWF multimers is directed largely by the sequence of the primary translation product rather than by endothelial- or megakaryocyte-specific machinery. More detailed analysis of the structure of recombinant vWF multimers is in progress.

Our analysis reveals similarities between the mature vWF subunit and the large vWF pro-polypeptide. These two related glycoproteins are both secreted after cleavage⁶, but possible functions of the pro-polypeptide in the circulation, or in linkage of the N-termini of dimers during multimer assembly, remain undefined. Site-directed mutagenesis and expression of the cDNA in cultured cells can now be used to identify features of pro-vWF that are important for its processing and biological activity.

This work was partly supported by NIH grants to R.I.H. and S.H.O. We thank Patricia Murtha and Monique Davies for technical help, and Dr Robert Wise for comments and discussion. S.H.O. is an Investigator of the Howard Hughes Medical Institute.

Received 2 May; accepted 8 August 1986.

- Legaz, M. E. *et al.* *J. biol. Chem.* **248**, 3946–3955 (1973).
- Shapiro, G. A. *et al.* *J. clin. Invest.* **52**, 2198–2210 (1973).
- Wagner, D. D. & Marder, V. J. *J. biol. Chem.* **258**, 2065–2067 (1983).
- Lynch, D. C. *et al.* *J. biol. Chem.* **258**, 12757–12760 (1983).
- Wagner, D. D. & Marder, V. J. *J. Cell Biol.* **99**, 2123–2130 (1984).
- Fay, P. J. *et al.* *Science* **232**, 995–998 (1986).
- Perlman, D. & Halvorson, H. O. *J. molec. Biol.* **167**, 391–409 (1983).
- Federici, A. B. *et al.* *J. clin. Invest.* **74**, 2049–2055 (1984).
- Sadler, J. E. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **82**, 6394–6398 (1985).
- Shelton-Inloes, B. B., Titani, K. & Sadler, J. E. *Biochemistry* **25**, 3164–3171 (1986).
- Counts, R. B., Paskell, S. L. & Elgee, S. K. *J. clin. Invest.* **62**, 702–709 (1978).
- Hessel, B. *et al.* *Thromb. Res.* **35**, 637–651 (1984).
- Ohmori, K. *et al.* *J. Cell Biol.* **95**, 632–640 (1982).
- Slayter, H. *et al.* *J. biol. Chem.* **260**, 8559–8563 (1985).
- Fowler, W. E. *et al.* *J. clin. Invest.* **76**, 1491–1500 (1985).
- Mosher, D. F. *Prog. Hemost. Thromb.* **5**, 111–151 (1980).
- Doolittle, R. F. *Scient. Am.* **245**, 126–130 (1981).
- Pierschbacher, M. D. & Ruoslahti, E. *Nature* **309**, 30–33 (1984).
- Pytela, R. *et al.* *Science* **231**, 1559–1562 (1986).
- Ruoslahti, E. & Pierschbacher, M. D. *Cell* **44**, 517–518 (1986).
- Kaufman, R. J. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **83**, 3136–3140 (1986).
- Goodman, R. H. *et al.* *J. biol. Chem.* **257**, 1156–1159 (1982).
- Vasicek, T. J. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **80**, 2127–2131 (1983).
- Bell, G. I. *et al.* *Nature* **304**, 368–371 (1983).
- Wiborg, O. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **81**, 1067–1069 (1984).
- Rosenfeld, M. G., Amara, S. G. & Evans, R. M. *Science* **225**, 1315–1320 (1984).
- Leiter, A. B. *et al.* *J. biol. Chem.* **260**, 13013–13017 (1985).
- Warren, T. G. & Shields, D. *Cell* **39**, 547–555 (1984).
- Bockenstedt, P., Greenberg, J. M. & Handin, R. I. *J. clin. Invest.* **77**, 743–749 (1986).
- Bockenstedt, P., McDonagh, J. & Handin, R. I. *J. clin. Invest.* **78**, 551–556 (1986).
- Lynch, D. C. *et al.* *Cell* **41**, 49–56 (1985).
- Ginsberg, D. *et al.* *Science* **228**, 1401–1406 (1985).
- Vervaeke, C. L. *et al.* *Nucleic Acids Res.* **13**, 4699–4717 (1985).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
- Biggin, M. D., Gibson, T. J. & Hong, G. F. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3963–3965 (1983).
- Poncz, M. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **79**, 4298–4302 (1982).
- Dale, R. M. K., McClure, B. A. & Houchins, J. P. *Plasmod. Res.* **13**, 31–40 (1985).
- Chen, E. Y. & Seeberg, P. H. *DN A* **4**, 165–170 (1985).
- Vieira, J. & Messing, J. *Gene* **19**, 259–268 (1982).
- Wong, G. G. *et al.* *Science* **228**, 810–815 (1985).
- Orkin, S. H. *et al.* *Molec. cell. Biol.* **5**, 762–767 (1985).
- Sandri-Goldin, R. M. *et al.* *Molec. cell. Biol.* **1**, 743–752 (1981).
- Lopata, M. A., Cleveland, D. W. & Sollner-Webb, B. *Nucleic Acids Res.* **12**, 5707–5717 (1984).
- Luthman, H. & Magnusson, G. *Nucleic Acids Res.* **11**, 1295–1308 (1983).
- Kaufman, R. J. *Proc. natn. Acad. Sci. U.S.A.* **82**, 689–693 (1985).
- Kaufman, R. J. *et al.* *Molec. cell. Biol.* **5**, 1750–1759 (1985).
- Laemmli, U. K. *Nature* **227**, 680–685 (1970).
- Bonthon, D. T. *et al.* *Nucleic Acids Res.* (in the press).

A nocturnal inhibitor of carboxylation in leaves

S. Gutteridge*§, M. A. J. Parry*, S. Burton*,
A. J. Keys*, A. Mudd*, J. Feeney†,
J. C. Servaites‡ & J. Pierce§

* Rothamsted Experimental Station, Harpenden,
Herts AL5 2JQ, UK

† National Institute for Medical Research, Mill Hill,
London NW7 1AA, UK

‡ Department of Biology, University of Dayton, Dayton,
Ohio 45469-0001, USA

§ Central Research and Development Department,
Experimental Station, E.I. DuPont Company, Wilmington,
Delaware 19898, USA

The diurnal variation in the activity of ribulose-1,5-bisphosphate carboxylase (RuBPCase), the major CO_2 -fixing enzyme in plants, has been shown to result from the influx and efflux of Mg^{2+} ions into and out of the chloroplast stroma. A recent re-examination of the phenomenon indicates that the inactivation of the enzyme, rather than being due to the efflux of Mg^{2+} , is correlated in some plant species with an increase in the concentration of an organic phosphate ester in the chloroplast in the dark¹⁻³. We have purified this potent inhibitor from dark-treated potato (*Solanum tuberosum*) leaves, and established that its structure is 2-carboxy-D-arabinitol-1-phosphate, a molecule that closely resembles an intermediate in the carboxylase reaction of RuBPCase.

The method used to isolate the inhibitor from dark-treated plant material exploited its affinity for the active enzyme. The carboxylase was extracted and purified in buffer containing CO_2 (HCO_3^-) and Mg^{2+} so that the inhibitor remained bound. The specific carboxylase activity of the enzyme at the first stages of purification was only 40% that of the enzyme purified from leaves during the day. This activity could be more than doubled by treatment with sulphate to release the inhibitor from the enzyme¹. Full details of the purification procedure are given in the legend to Table 1. Figure 1 shows the inhibition of purified wheat RuBPCase as a function of the total phosphate content of the purified inhibitor at various dilutions. Maximum inhibition was seen at inhibitor concentrations stoichiometric with those of the enzyme's active site.

There are a number of phosphate esters synthesized in the chloroplast⁴ (for example, 6-phosphogluconate) that may act as potent inhibitors of the carboxylase. The purified inhibitor was compared to such compounds by thin layer chromatography on cellulose powder and in several solvent systems. The R_f values of the natural inhibitor and its dephosphorylated product indicated that it was unlike any of the phosphorylated sugars or sugar alcohols usually found in plants.

A sample containing 2.5 μmol of the inhibitor (based on total phosphate) was dissolved in 0.4 ml of D_2O ($pD \sim 3.0$) and the solution was examined by proton NMR spectroscopy using a

Table 1 Inhibitor purification from dark-harvested potato leaves

Fraction	Weight of dry residue (mg)	Total phosphorus ($\mu\text{g atom}$)
Extract in 80% methanol	1,500	11.9
Sephadex G-10 column	92	10.6
Polar fraction from C_{18} Sep-Pak	16	7.6
QAE-A25 column	1.7	6.1

To obtain enough material for detailed analysis, 1 kg of potato leaves were collected early in the morning and the soluble proteins extracted at pH 8.0 in a buffer containing 20 mM MgCl_2 and 20 mM NaHCO_3 . RuBPCase with the inhibitor bound was isolated from the plant debris as a precipitate using polyethylene glycol. The precipitate was redissolved in a minimum volume of the activation buffer, 0.1 M bicine pH 8.0 containing 20 mM MgCl_2 and 20 mM NaHCO_3 . Further purification was achieved by chromatography on an S400 (Pharmacia) gel filtration column. The inhibitor was released from the enzyme by addition of methanol to give a final concentration of 80% v/v. Precipitated protein, 1.3 g, was removed by centrifugation and re-extracted with further additions of 80% methanol. The combined extracts were concentrated and desalted into a volatile buffer (10 mM NH_4HCO_3) using a Sephadex G-10 column and those fractions exhibiting the strongest inhibition of wheat RuBPCase were evaporated to dryness. This step removed MgCl_2 introduced with the buffer used in extraction and purification of the potato enzyme. The residue was dissolved in distilled water, adjusted to pH 4.0 with formic acid, and further purified by passage through a C_{18} Sep-Pak column (Waters). The final treatment was consecutive ion-exchange chromatography using anionic and cationic columns⁹. The resulting compound, dried to constant weight, was suitable for spectrophotometric analysis.

Bruker 500 MHz instrument. The spectrum consists of a number of multiplets arising from only six protons associated with four C centres. Selective irradiation spin decoupling experiments indicate that three of these centres are adjacent to one another. Consideration of the spin coupling constants and the chemical shifts for these protons suggests a structure of $\text{HOCH}_2\text{-CH(OH)-CH(OH)-}$. The remaining multiplets correspond to a pair of methylene protons that are not coupled to the other protons in the molecule and have chemical shifts and additional splittings ($J_{\text{P-}^1\text{H}} = 5 \text{ Hz}$) characteristic of a $-\text{CH}_2\text{OPO}_3^{2-}$ group⁵. Thus the inhibitor has the general structure $\text{CH}_2(\text{OH})\text{-CH(OH)-CH(OH)-X-CH}_2\text{OPO}_3^{2-}$. The NMR spectrum is pH dependent, consistent with the presence of a carboxylic acid group as a branch to the carbon chain, and at low pH cyclization to a lactone ring can occur. X was therefore considered to be a quaternary carbon with a carboxylic group, that is $>\text{C}(\text{OH})\text{-CO}_2^-$. This suggested that the compound might be similar to the potent carboxypentitol inhibitors 2-carboxyarabinitol-1,5-bisphosphate and 2-carboxyribitol-1,5-bisphosphate (2CABP and 2CRBP, respectively) but with a single phosphate group at the 1 position.

The confirmation that the natural inhibitor must be a 2-carboxy pentitol-1-phosphate was achieved by first dephos-

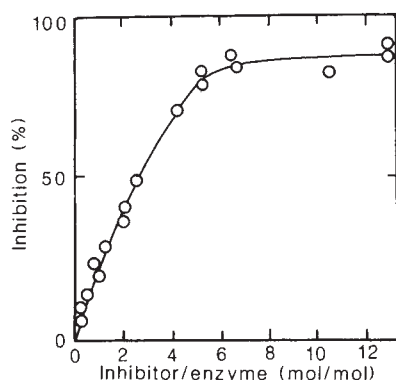


Fig. 1 Inhibition of carboxylase activity of purified wheat RuBPCase⁸ by the inhibitor from potato leaf. The carboxylase was activated (2 mg ml⁻¹) in 0.1 M bicine, pH 8.2 at 40 °C for 40 min in the presence of 10 mM NaHCO_3 and 20 mM MgCl_2 . Samples of this solution (20 μl) were mixed with 20 μl of a solution containing the inhibitor. After at least 2 min, 20 μl of the mixture was assayed for carboxylase activity at 25 °C. The inhibitor concentration is based on the phosphate content of the solutions.

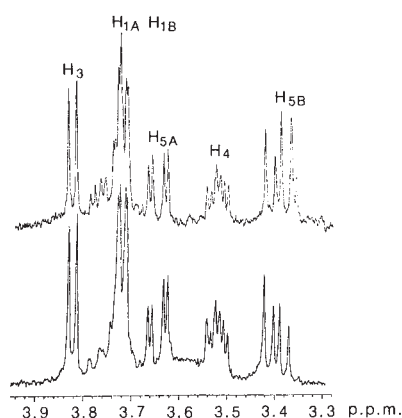


Fig. 2 The proton NMR spectrum at 500 MHz of 2-carboxyarabinitol-1-phosphate. *a*, The natural compound, isolated by dissociation from the active site of RuBPCase and then purified as outlined in the legend to Table 1. *b*, The spectrum of the monophosphate synthesized from 2-carboxyarabinitol-bisphosphate by partial hydrolysis with alkaline phosphatase (S.G. and M.A.J.P., unpublished data). The chemical shifts are given in Table 2. The *pD* of the samples was 8.0 and hence the spectra are for the open chain form.

phorylating the compound with alkaline phosphatase and then producing the trimethyl silylated derivative for examination by gas chromatography mass spectrometry. One of the major silylated products has an ion ($m/z = 613$, $M-15^+$) that is consistent with a fully trimethyl silylated derivative of carboxypentitol. A second ion ($m/z = 466$, M^+) was also observed corresponding to the trimethyl silylated lactone form of a carboxypentitol. The same molecular ions were produced when 2-carboxyarabinitol was synthesized from 2CABP by phosphatase treatment and silylated under the same conditions.

There are four epimers that satisfy the general structure described above. The natural compound could be 2-carboxyarabino- or 2-carboxyribo- stereoisomers (epimeric about C_2) or 2-carboxylyxitol or 2-carboxyxylitol monophosphates (epimeric about C_3). To distinguish between these possibilities, the bisphosphates were synthesized as previously described^{6,7}, the phosphate groups were removed and the proton spectra of the resulting acids compared to the dephosphorylated natural inhibitor. It is quite clear from the chemical shifts of the protons (Table 2) that the compound is identical to 2-carboxy-D-arabinitol; therefore, the natural phosphate ester is 2-carboxy-D-arabinitol-1-phosphate. Its proton spectrum is shown in Fig. 2 along with the spectrum of the monophosphate synthesized from 2CABP.

2-carboxyarabinitol-1-phosphate has many features enabling

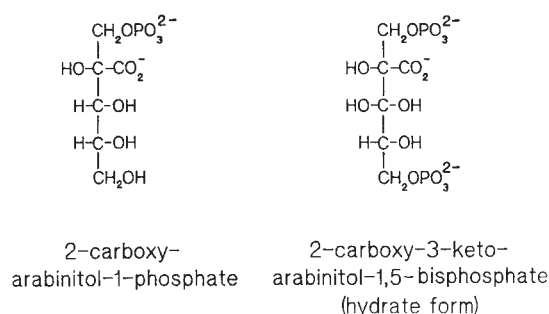


Fig. 3 The structure of the natural inhibitor 2-carboxyarabinitol-1-phosphate compared to a reaction intermediate of the carboxylase reaction of RuBPCase.

it to act as a potent and yet reversible inhibitor of RuBPCase. It closely resembles the 2-carboxy-3-ketoarabinitol-1,5-bisphosphate intermediate of the carboxylase reaction (see Fig. 3). However, unlike 2CABP the absence of a second phosphate group on the primary alcohol group at the five position ensures that it does not become bound irreversibly to the active site. The response of RuBPCase to effectors, particularly transition state analogues, in terms of activation and activity *in vitro* have been exhaustively studied (see ref. 4 for a review), but this is the first report of a compound with this type of structure functioning in this way in the chloroplast. The function of the transition state analogue as an inhibitor of RuBPCase *in vivo* is unknown. During normal photosynthesis, the enzyme is activated by carbamylation with CO_2 thus forming the binding site for the essential Mg^{2+} ion. The substrate, ribulose bisphosphate, binds to the active site and the close interaction with the metal initiates the catalytic process⁵. However, if the substrate binds to the active site before the reaction with CO_2 and Mg^{2+} it very effectively blocks the activation process. An inhibitor that can dissociate from the active site and yet is preferentially bound to the activated form of the enzyme, might serve to stabilize this conformation during periods of reduced light or in darkness to ensure it is primed for rapid substrate turnover when photosynthesis recommences. Certainly, the rates of appearance and disappearance of the inhibitor are consistent with such a function.

The metabolic precursor of the inhibitor has not yet been determined. The inhibitor must be a major component of dark-induced metabolism because the concentration of RuBPCase active sites in the stroma is at least 3 mM and therefore 2-carboxyarabinitol-1-phosphate must accumulate to these concentrations to cause significant inhibition. The inhibition of the enzyme at night has not been detected in all plants³. Until the biosynthetic enzymes that regulate the inhibitor concentration

Table 2 Chemical shifts* in p.p.m. relative to dioxane of the natural inhibitor, the dephosphorylated derivative and synthetic compounds

Compound	C-1		C-3	C-4	C-5	
	A	B			A	B
Natural inhibitor	3.73	3.71	3.82	3.52	3.64	3.40
2-Carboxyarabinitol-1-phosphate (open chain)	3.73	3.71	3.82	3.52	3.64	3.40
2-Carboxyarabinitol-1-phosphate (lactone)	3.82	3.72	4.19	4.24	3.80	3.58
Dephosphorylated inhibitor	3.59	3.53	3.70	3.51	3.63	3.40
2-Carboxyarabinitol	3.57	3.53	3.70	3.51	3.63	3.40
2-Carboxyribitol	(3.49–3.55)		(3.49–3.55)	3.37	(3.49–3.55)	
2-Carboxyxylitol	3.58	3.53	(3.54–3.58)	(3.54–3.58)	3.35	3.33
2-Carboxylyxitol	3.60	3.48	3.52	3.70	3.42	3.38

* Overlapping multiplets are shown as a range of p.p.m. values.

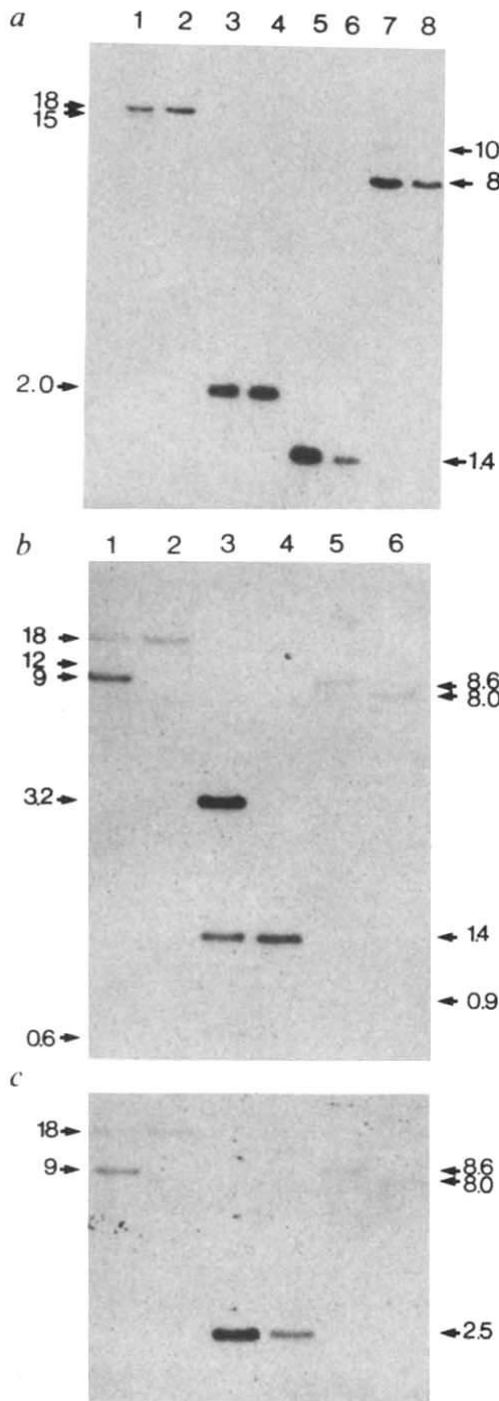


Fig. 2 DNA rearrangement at the *c-myc* locus in W74 HCC and W93 HCC. *a*, Hybridization of W74 HCC and W20 liver DNA with *c-myc* exon 1; *b*, hybridization of W93 tumour and non-tumour liver DNA with *c-myc* exon 1; *c*, subsequent hybridization of the same filter with *c-myc* exon 2. *a*, Cellular DNA (30 µg) from W74 HCC (lanes 1, 3, 5, 7) or W20 normal liver (lanes 2, 4, 6, 8) was digested with: 1, 2, *Eco*RI; 3, 4, *Bam*HI; 5, 6, *Bgl*II; 7, 8, *Hind*III. *b, c*, DNA from W93 tumour (lanes 1, 3, 5) and non-tumour (lanes 2, 4, 6) liver samples was digested with: 1, 2, *Eco*RI; 3, 4, *Bgl*II; 5, 6, *Hind*III. Southern blotting was performed as described²¹. For probes and conditions of hybridization see Fig. 1 legend. The restriction pattern of DNA from the non-tumorous part of W74 liver was identical to that of W20 liver DNA and thus is not shown. The sizes of the *c-myc*-containing fragments are given in kb. DNA size markers were bacteriophage λ DNA digested with *Hind*III and double digested with *Eco*RI and *Hind*III. The blots were exposed to X-ray films for 5 days. The filters were reprobbed with ³²P-labelled pBR322 DNA to ensure that all bands were *myc*-specific.

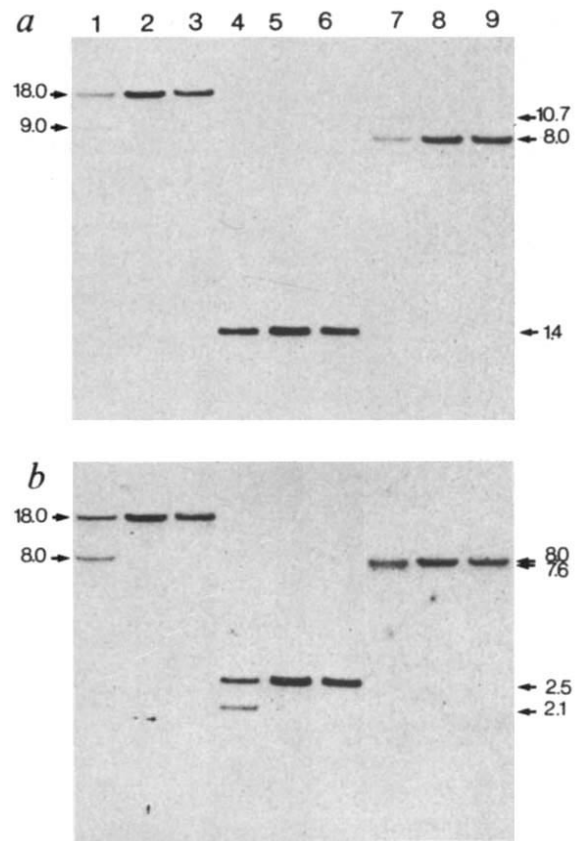


Fig. 3 DNA rearrangement at the *c-myc* locus in W64 HCC. *a*, Southern blotting analysis of genomic DNA from W64 HCC (lanes 1, 4, 7), non-tumorous parts of W64 liver (lanes 2, 5, 8) and W20 normal liver (lanes 3, 6, 9) probed with *myc* exon 1. Lanes 1-3 are digestions with *Eco*RI, lanes 4-6 with *Bgl*II, and lanes 7-9 with *Hind*III (Figs 1 and 2 legends). *b*, The same filter was rehybridized with *myc* exon 2. Hybridization of the filter with a probe derived from pBR322 did not reveal any band. Sizes in kb.

tive of woodchuck *c-myc* exon 1 and human *c-myc* exons 2 and 3. The results obtained with *myc* exon 3 probe (not shown) were identical to those of *myc* exon 2. Densitometric analysis indicated that *c-myc* RNA levels were elevated at least 5-, 10- and 50-fold in W74, W93 and W64 HCC, respectively compared with the corresponding adjacent liver tissues from the same animal. The level of *c-myc* transcription was below the threshold of detection in normal adult woodchuck liver. The same probes revealed abundant 2.3-kb *c-myc* transcripts in the RNA of human HL60 cells (Fig. 1*d*). However, hybridization of HL60 RNA with a *c-myc* exon 1 probe of woodchuck origin gave a considerably weaker signal than with the human *c-myc* exon 2 probe.

In contrast with the apparently normal 2.3-kb *c-myc* RNA found in the non-tumorous parts of the livers, the tumours produced *c-myc* transcripts of different sizes. W74 HCC produced a major 5.6-kb *c-myc* RNA which annealed to all three *c-myc* exons (Fig. 1*a*). The *c-myc* transcripts in W93 HCC also hybridized with all three *c-myc* exons (Fig. 1*b*); they migrated as a heterogeneous population of 2.3–2.5 kb, slightly larger than normal. This could reflect a shift in *c-myc* promoter usage from P2 to P1, as seen in Burkitt's lymphomas^{1,2} and murine plasma cell tumours³. W64 HCC had a major truncated 2.0-kb *c-myc* messenger RNA, and two minor mRNAs of 5.8 and 4.7 kb (Fig. 1*c*). These RNAs annealed to human *c-myc* exons 2 and 3, but not to woodchuck *c-myc* exon 1 (Fig. 1*c*), indicating that the normal transcriptional unit was disrupted in the tumour as has been reported for B-lymphoid tumours^{4–7}.

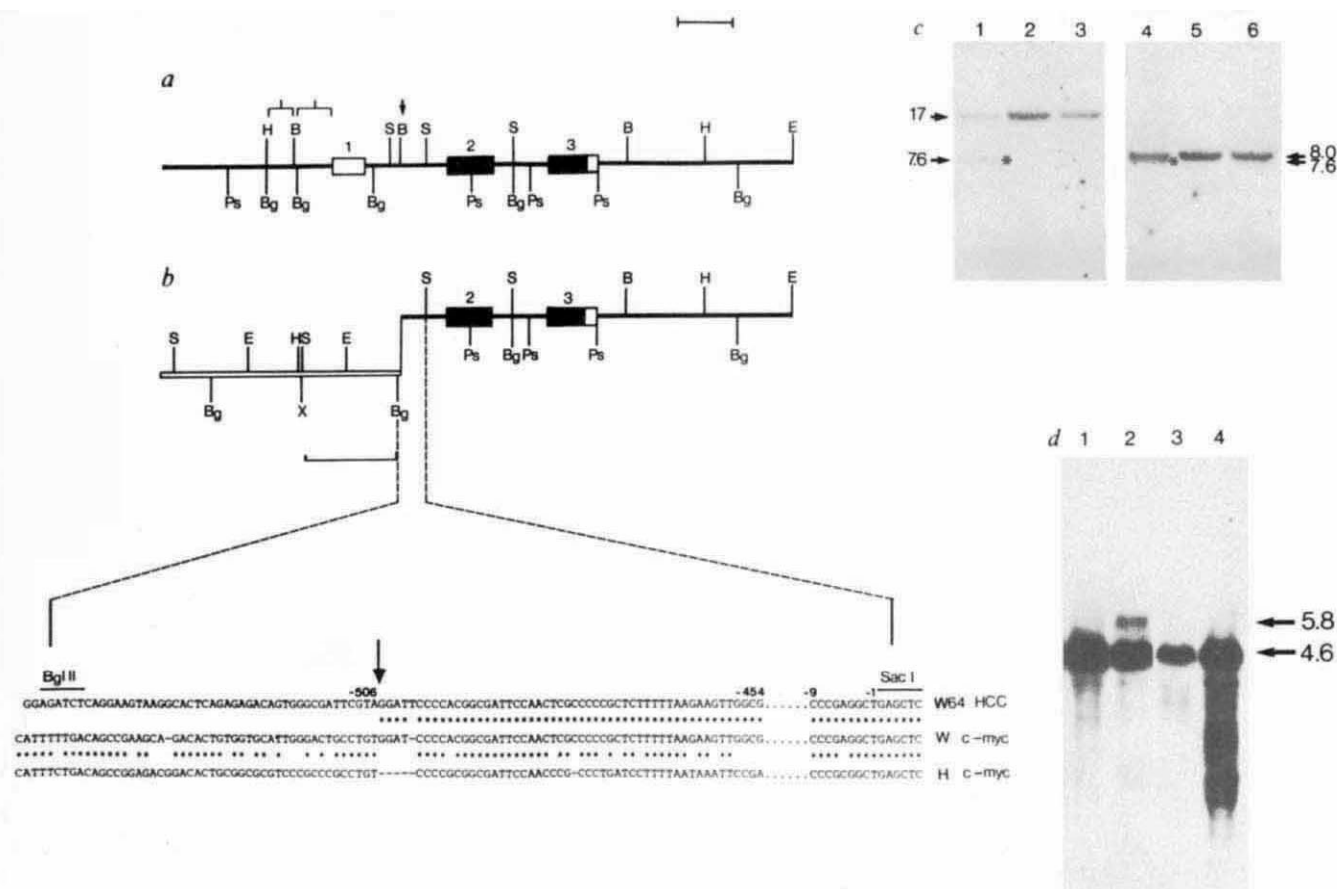


Fig. 4 *a*, Schematic representation of the normal woodchuck *c-myc* allele. Brackets, the recombination break points for the W74 HCC and W93 HCC as deduced from genomic blot studies; arrow, the precise breakpoint for W64 HCC (determined by molecular cloning and sequencing). The normal *c-myc* gene was mapped from restriction endonuclease digests of cloned DNA from W64 HCC as described (Fig. 1 legend). Scale bar, 1 kb. *b*, Mutant *c-myc* allele of W64 and sequence analysis of the break-point region. Unshaded bar, incoming cellular sequences; small bar, probe s/bg; lower panel, enlargement of the region between *Bgl*II and *Sac*I sites. The sequence of the mutant *c-myc* allele in the break-point area (W64 HCC) was aligned with the corresponding sequences of the normal *c-myc* intron 1 from woodchuck (W *c-myc*) and man (H *c-myc*). Asterisks, homologous nucleotides; arrow, breakpoint for W64 HCC. *c*, Southern blotting analysis of *Hind*III-digested DNA from: lanes 1, 4 W64 HCC; 2, 5, W64 non-tumorous part of liver; 3, 6, W20 normal liver. Hybridization was performed as described (Fig. 2 legend). Lanes 1–3 were probed with the 1.7 kb *Sac*I–*Bgl*II fragment located immediately upstream of the break point (small bar, Fig. 4*b*); lanes 4–6 were probed with *c-myc* exon 2. Asterisks denote the tumour-specific fragment of 2.6-kb which also hybridizes with exon 2 (see Fig. 3). *d*, Northern blotting analysis of total cellular RNA. Lane 1, non-tumorous part and lane 2, tumorous part of W64 liver; lane 3, normal liver W101; lane 4, normal liver, W20. Hybridization to the s/bg probe was as described in Fig. 1 legend.

Methods. Total cellular DNA from W64 HCC was partially digested with *Sau*3A. 10–20-kb fragments were selected and ligated to the *Bam*HI-digested arms of the vector λ -L47.1. A library (2×10^6 phages) was screened with the human *c-myc* exons 2 and 3 (see Fig. 1*e*), using the Benton and Davies *in situ* hybridization procedure²⁹. Eleven independent clones were obtained and analysed by restriction mapping and hybridization. Five clones covered the entire normal *c-myc* gene as represented in Fig. 4*a*. Three clones retained *c-myc* exons 2 and 3 but contained non-*myc* sequences upstream of exon 2. The restriction map of the mutant *c-myc* allele was established from the DNA of these three clones (*b*), and is consistent with the data obtained by Southern blotting (Fig. 3). The 2.3-kb *Sac*I fragment of the mutant allele, encompassing the break-point region was subcloned in a plasmid vector and further cloned into the M13 phage vector mp8, using the method of Deininger *et al.*³⁰. The 0.7-kb *Sac*I fragment from the normal *c-myc* intron 1 was directly subcloned in M13 tg130. The sequencing reactions were performed as described by Sanger *et al.*³¹.

Cellular DNA from W74, W93 and W64 HCC was analysed for rearrangement or amplification of the *c-myc* locus. Figure 2 compares Southern blots of genomic DNA from W74 and W93 HCC with W20 normal liver and with the non-tumorous part of W93 liver, respectively. Additional *c-myc* bands were revealed in these HCC, indicating that a *c-myc* allele has been rearranged in the tumour cells. The novel DNA fragments were not due to polymorphism in the length of restriction fragments, since identical normal *c-myc* bands were detected in DNA samples from the non-tumorous parts of livers W74 and W93 and from W20 (Fig. 2) and 16 other woodchuck livers (not shown). Using woodchuck *myc* exon 1 as a probe, tumour-specific fragments were observed in *Eco*RI and *Hind*III digests

of W74 HCC DNA, but not in *Bam*HI and *Bgl*II digests (Fig. 2*a*). Hybridization with *myc* exons 2 and 3 revealed the same novel *Eco*RI and *Hind*III fragments and normal 2.5-kb and 4.1-kb *Bgl*II bands (not shown). Comparison of genomic blot analysis of W74 HCC DNA with the restriction map of a cloned normal woodchuck *c-myc* allele suggests a recombination event outside, and probably upstream of, the *c-myc* gene (Fig. 4). In W93 HCC, novel 9-kb *Eco*RI and 8.6-kb *Hind*III fragments hybridized with *myc* exons 1 and 2 probes, and the faint 12-kb *Eco*RI and 0.9-kb *Hind*III fragments with *myc* exon 1 probe only (Fig. 2*b,c*). Tumour-specific *Bgl*II bands were revealed only after hybridization with *myc* exon 1. This suggests a complex recombination event, next to the 5' end of the *c-myc* gene

(Fig. 4). Southern blot analysis of DNA isolated from different parts of W93 HCC revealed that rearrangements of the *c-myc* locus were restricted to one part of the tumour; this tumour was thus oligoclonal. Enhanced expression of *c-myc* was found only in the tumour samples in which the *c-myc* gene was altered. In contrast, W64 and W74 HCC appeared monoclonal after several blot experiments (results not shown).

W64 HCC also showed rearrangement at the *c-myc* locus (Fig. 3). The probes for *myc* exons 1 and 2 annealed with the same normal *EcoRI* fragment, but with distinct tumour-specific bands. This was also observed with *HindIII* digests. The results obtained using a *myc* exon 3 probe (not shown) were identical to those for *myc* exon 2. Digestion with *BglIII* revealed a novel band specific to *myc* exon 2 (Fig. 3b). The *c-myc* DNA rearrangement in W64 HCC thus takes place within the first intron or the second exon. The faint signals observed for tumour-specific fragments annealing to woodchuck *myc* exon 1 may indicate a partial loss of rearranged *myc* exon 1. We cloned the rearranged *c-myc* allele from a λ phage library of W64 HCC DNA, and located the recombination site in *c-myc* intron 1 by restriction mapping (Fig. 4b). A 2.3-kb *SacI* fragment, encompassing the joining region between *c-myc* and partner sequences, was subcloned and sequenced. We compared this sequence with those of normal *c-myc* alleles from woodchuck (cloned from W64 HCC DNA) and man⁸. Figure 4b (lower panel) shows that the breakpoint occurs 505 nucleotides upstream of the *SacI* site nearest to the 5' boundary of *myc* exon 2, which is found in both species. Thus the mutant *c-myc* allele retains the cryptic promoters and initiation sites of intron 1, which are used for the transcription of truncated *c-myc* genes in human Burkitt's lymphoma⁹, acute B-cell leukaemia¹⁰, and murine plasmacytomas^{5,7}. In such cases, as in the woodchuck tumour, recombination *myc* exon 1 or intron 1 results in the production of novel RNA species that are shorter than the normal *c-myc* transcripts¹¹.

To improve our understanding of the mechanisms of *c-myc* activation in this tumour, we characterized the invading sequence next to the *c-myc* coding exons. Hybridization studies of the cloned mutant *c-myc* allele from W64 HCC, using pWHV-1 as a probe¹², showed that there were no WHV sequences in the 5-kb regions flanking the 5' and 3' sides of the *c-myc* gene (not shown). A 1.7-kb *SacI*-*BglIII* fragment (denoted s/bg in Fig. 4b) immediately next to the rearranged *c-myc* gene was used to probe Southern blots of W64 and W20 DNA, and revealed a unique 17-kb *HindIII* band in normal woodchuck liver DNA. Another tumour-specific fragment of 7.6 kb in W64 HCC DNA annealed with both s/bg and *myc* exon 2 (Fig. 4c). Further blot studies showed that the s/bg probe anneals to a unique sequence in the human genome (results not shown). The same probe revealed an abundant 4.6-kb transcript in W64 and normal W101 and W20 livers (Fig. 4d) and other woodchuck

livers (not shown). In W64 HCC, it also annealed to a 5.8-kb transcript migrating at the same position as the longer minor *c-myc* RNA which could therefore be a co-transcript of s/bg and *c-myc* exons 2 and 3. A computer-assisted search of the Los Alamos database failed to identify the s/bg sequence with any known cellular gene, including the immunoglobulin locus and the T-cell receptor gene.

Activation of *c-myc* by different mechanisms has been implicated in a variety of haematopoietic malignancies and carcinomas^{13,14}. In some solid tumours, *c-myc* amplification^{15,16} or rearrangement in a few cases¹⁶ leads to an increased level of expression, but no major alteration of the *c-myc* locus has been reported in rodent liver tumours induced by chemical carcinogens¹⁷. In B and T lymphoid tumours, *c-myc* expression is deregulated by a structural alteration of the gene, either chromosomal translocation^{5,10} or insertional mutagenesis with non-acute oncogenic retroviruses^{18,19}. The DNA rearrangements and enhanced expression of *c-myc* in viral-associated HCC reported here seem quite similar to those found in B and T cell neoplasias. It has been proposed that integration of hepatitis viruses in the cellular DNA of their natural hosts, observed almost invariably in liver tumour cells, is critical to HCC development²⁰. Indeed, Southern blot and cloning data revealed the presence of integrated WHV sequences in the three woodchuck HCC examined here (ref. 21 and our unpublished results), but our data exclude *c-myc* activation by insertional mutagenesis in W64 HCC. We have no evidence so far for viral insertion near *c-myc* in the other two tumours. The recombination of a truncated *c-myc* gene with another cellular gene, described here for W64 HCC, is reminiscent of the reciprocal translocations of *c-myc* and immunoglobulin or T-cell receptor genes which occur in human Burkitt's lymphoma, B- and T-cell acute lymphoid leukaemia and mouse plasmacytoma. Furthermore, it has recently been pointed out that in Burkitt's lymphoma and mouse plasmacytoma, *c-myc* alterations occur after a long pre-neoplastic period of chronic cell stimulation, provoked in some cases by Epstein-Barr or Abelson virus²². A similar situation could result from chronic hepatitis virus infection in woodchuck and man, favouring chromosomal breaks²³ and finally leading to the development of HCC.

We thank Dr S. Aaronson, S. Alonso, E. Chang, T. Curran, K. Marcu, C. Sherr and D. Stehelin for providing *onc*-specific and actin clones, C. Pichoud and C. Jacquet for maintenance of woodchucks and surgical biopsies, M. Hadchouel for histological examination of tumour tissue, J. Robert-Lézènes for helpful discussions, J. Davies for a critical reading of the manuscript and L. M. Da and A. Cova for typing the manuscript. T.M. is supported by the Commission of the European Community. This work was supported by NIH grant CA 97300-02, the Economic European Community and the Association pour la Recherche contre le Cancer.

Received 21 April; accepted 8 September 1986.

- Hollis, G. F. *et al.* *Nature* **307**, 752-755 (1984).
- Taub, R. *et al.* *Cell* **36**, 339-348 (1984).
- Yang, J., Bauer, S. R., Muchinski, J. F. & Marcu, K. B. *EMBO J.* **4**, 1441-1447 (1985).
- Shen-Ong, G., Keath, E., Piccoli, S. & Cole, M. *Cell* **31**, 443-452 (1982).
- Stanton, L. W., Watt, R. & Marcu, K. B. *Nature* **303**, 401-406 (1983).
- Adams, J. M., Gerondakis, S., Webb, E., Corcoran, L. & Cory, S. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1982-1986 (1983).
- Rushdi, A. *et al.* *Science* **222**, 390-393 (1983).
- Bernard, O., Cory, S., Gerondakis, S., Webb, E. & Adams, J. *EMBO J.* **2**, 2375-2383 (1983).
- Hayday, A. C. *et al.* *Nature* **307**, 334-340 (1984).
- Care, A. *et al.* *EMBO J.* **5**, 905-911 (1986).
- Perry, R. P. *Cell* **33**, 647-649 (1983).
- Ogston, C., Jonak, G., Rogler, C., Astrin, S. & Summers, J. *Cell* **29**, 385-394 (1982).
- Varmus, H. E. *A. Rev. Genet.* **18**, 553-612 (1984).
- Weinberg, R. A. *Science* **230**, 770-776 (1985).
- Yokota, J., Tsunetsugu-Yokota, Y., Battifora, H., Le Fevre, C. & Cline, M. J. *Science* **231**, 261-265 (1986).
- Escot, C. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **83**, 4834-4838 (1986).
- Yaswen, P., Goyette, M., Shank, P. R. & Fausto, N. *Molec. cell Biol.* **5**, 780-786 (1985).
- Hayward, W. S., Neel, B. G. & Astrin, S. M. *Nature* **290**, 475-480 (1981).
- Corcoran, L. M., Adams, J. M., Dunn, A. R. & Cory, S. *Cell* **37**, 113-122 (1984).
- Dejean, A., Bougueleret, L., Grzeschick, K. H. & Tiollais, P. *Nature* **322**, 70-72 (1986).
- Möröy, T., Etienne, J., Trépo, C., Tiollais, P. & Buendia, M. A. *EMBO J.* **4**, 1507-1514 (1985).
- Klein, G. & Klein, E. *Nature* **315**, 190-195 (1985).
- Rogler, C. *et al.* *Science* **230**, 319-322 (1985).
- Collins, S., Gallo, R. & Gallagher, R. *Nature* **270**, 347-349 (1977).
- Collins, S. & Groudine, M. *Nature* **298**, 679-681 (1982).
- Thomas, P. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5201-5205 (1980).
- Alonso, S., Minty, A., Bourlet, Y. & Buckingham, M. *J. molec. Evol.* **23**, 11-22 (1986).
- Stanton, L. W., Fahrlauder, P. D., Tesser, P. M. & Marcu, K. B. *Nature* **310**, 423-425 (1984).
- Benton, W. D. & Davies, R. W. *Science* **196**, 180-182 (1977).
- Deininger, P. L. *Analyt. Biochem.* **129**, 216-223 (1983).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).

The role of nucleotide sequences in splice site selection in eukaryotic pre-messenger RNA

L. P. Eperon, J. P. Estibeiro & I. C. Eperon

Department of Biochemistry, University of Leicester, Leicester LE1 7RH, UK

Alternative splicing of eukaryotic messenger RNA precursors is now known to be of widespread importance in generating multiple transcripts from a single gene. This phenomenon has emphasized the problem of the way in which splice sites are selected; recent studies have discussed the role of secondary structure¹ or affinity and spatial relationships² in this selection. Splice site sequences vary widely, although a loose consensus has been derived for the 9 bases around the 5' splice site and for a longer region around the 3' splice site³. Mutagenesis experiments have defined the sequences essential for a potential 5' splice site^{4,5,6}, but, except for some experiments with the *Ela* gene of adenovirus^{5,6}, these experiments have not examined 5' splice site sequences for features responsible for site preference where alternative splicing sites exist. Such tests require a choice of site: an appropriate reference site and a constant position at which test sites are introduced. We have begun a series of experiments designed to show whether splice site sequences can be ranked in a hierarchy of preferential use. Here we show that the archetypal consensus sequence is used efficiently, and characterize the cryptic sites of β -globin: sequences alone can explain why these sites are not normally used. We also show with the *Ela* gene of adenovirus, a simple example of alternative splicing, that one of the two 5' splice sites used by this gene is intrinsically stronger. We also demonstrate that tandem repeats and secondary structure influence the choice of sites *in vivo*. We discuss the mechanism of splice site selection.

To define a hierarchy of splice site efficiency, we calibrated sequences by competition against a constant reference site. An assay was devised involving choice between alternative splice sites contained in short oligonucleotides and the 5' splice site of the second intervening sequence (IVS-2) of the rabbit β -globin gene (Fig. 1a). The splicing junction to be assessed was always 25 nucleotides upstream of the authentic site. The sequences tested (Table 1) vary at the 5' splice junction in the last three nucleotides of the exon and first six of the intron, the region defined by the consensus sequence. The 'cons' oligonucleotide is composed of the most frequently-used bases at each position in the consensus³; 'glo' is a replica of the reference site, a necessary control for the effect of position; the 'cry' oligonucleotides contain sequences corresponding to the cryptic β -globin sites which are revealed only when the 'GT' at the beginning of

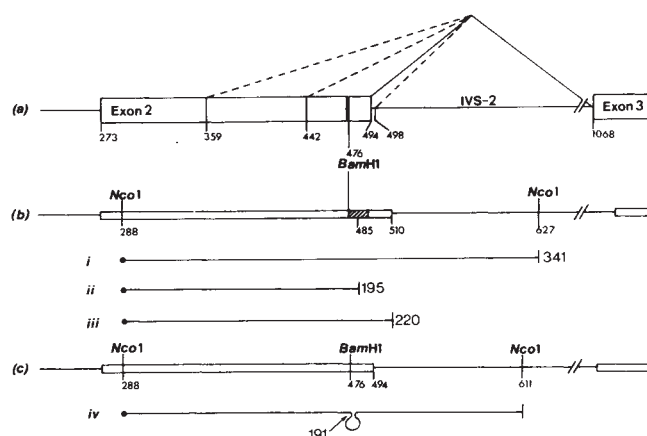


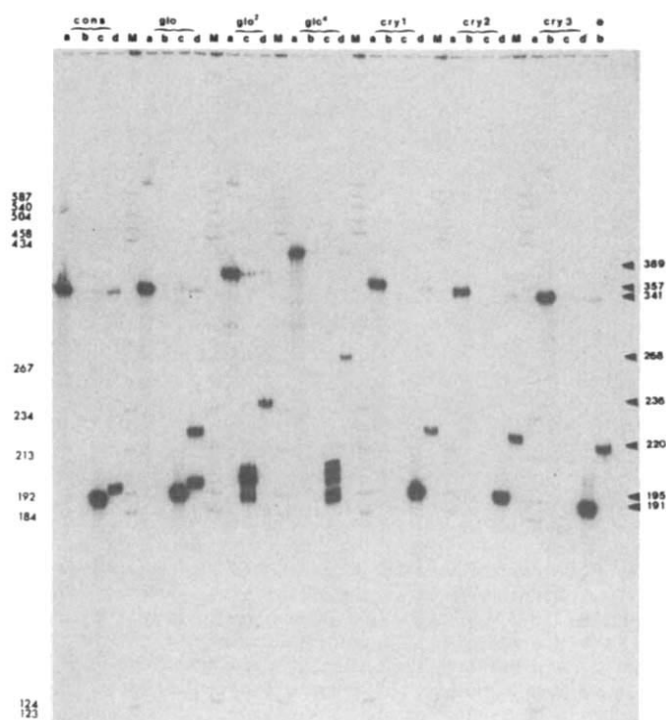
Fig. 1 a, Splicing of the second intervening sequence (IVS-2) of rabbit β -globin. The boxes represent exons. Solid lines between exons, authentic splicing pattern; dashed lines, splicing to cryptic sites when the 5' splice site of IVS-2 is mutated. b, explanation of the RNA mapping strategy; c, use of *in vitro* RNA corresponding to β -globin gene sequences. a, The position of the *Bam*HI site into which 16-base oligonucleotide pairs (see Table 1) were cloned is shown. The 'cons' and 'cry' paired oligonucleotides were first inserted into the *Bam*HI site of a *Bam*HI-*Bgl*II fragment of the β -globin gene which had been cloned into an M13 vector. The *Bam*HI-*Eco*RI oligonucleotide-containing fragment was used to create the appropriate derivative of p β SV via three-way ligations. The 'glo' oligonucleotides were inserted as a monomer, dimer (glo²) and a tetramer (glo⁴) directly into p β SV and the recombinants screened using the 'microprep' method. This involved growing 96 small cultures in a microtitre plate and preparing the DNA by a modification of the alkaline lysis 'miniprep' method¹¹. The products were analysed by restriction enzyme digestion. All the insertions into p β SV were sequenced using the dideoxy method¹² either directly on the double stranded plasmid or via recloning into M13 vectors. b, An inserted oligonucleotide is shown by a shaded box; the *Bam*HI site is restored only at the 5' end of the oligonucleotide. (i), Undigested, full length probe, an *Nco*I restriction fragment, 3' end-labelled by a two nucleotide extension at the staggered end using dCTP and [α -³²P]dATP and Klenow polymerase (indicated by a spot). This fragment was separated from other *Nco*I fragments on a 4% or 6% preparative sequencing gel containing 20% formamide. (ii), The product after hybridization of the probe to an RNA transcript which has used the new 5' splice site, and digestion with *S*₁ nuclease. (iii), The digested probe when the reference site was used for splicing. c, A *Pst*I-*Bgl*II fragment was cloned into pSP64, and RNA was transcribed *in vitro* from this. (iv), Hybridization of a probe made from β -globin with oligonucleotide inserts to the RNA transcribed *in vitro*. Digestion with *S*₁ nuclease would give a 191 nucleotide fragment. The RNA made *in vitro* is a valuable size marker.

Table 1 5' Splice site efficiencies

Name	Sequence	Test site (%)	Reference site (%)	Details
glo	GATCCCAGG GTGAGTC	64	36	IVS-2 β -globin 5' site From the optimal consensus ³
cons	GATCCCCAG GTAAGTC	100	0	
cry1(498)	GATCCCTGA GTTTGGC	0	100	β -globin cryptics
cry2(442)	GATCCCTAA GCTGAGC	0	100	
cry3(359)	GATCCCAAG GTGAAGC	0	100	
glo ²		0	100	
glo ⁴		0	100	Dimeric IVS-2 β -globin; The 5'-most oligonucleotide is in the complementary orientation. Tetrameric insert; the three 5'-most oligonucleotides are in the complementary orientation.
E1a ₉₇₃	GATCCCAGG GTGAGGC	0	100	Alternative E1a sites.
E1a ₁₁₁₁	GATCCCAAG GTAAGTC	63	37	
E1a ₂₁₁₁		100	0	
E1a ₁₁₁₁ ^{comp}		0	100	

2 correctly oriented oligonucleotides. The 5'-most site is used.
1 complementary sense oligonucleotide.

Fig. 2 Determination of the splice sites chosen using S_1 nuclease, for β -globin genes containing the following test sequences: the consensus (based on observed frequencies at each position³: 43% C, 64% A, 73% G, 100% G, 100% T, 62% A, 68% A, 84% G, 63% T), β -globin IVS-2 (glo), IVS-2 dimer (glo²), IVS-2 tetramer (glo⁴) and the β -globin cryptic sites (cry 1, 2, 3). Total RNA was extracted from HeLa cells with LiCl₂-urea¹³ 48 h after calcium phosphate transfection. RNA was hybridized to the *Nco*I double-stranded probe at 49 °C for 2 h. Standard conditions were used for S_1 nuclease digestions¹⁴. Samples were separated on a sequencing gel¹⁵ comprising 6% acrylamide, 20% formamide. Size markers (M) were [³²P]-phosphorylated *Hae*III-cut pBR322. Their sizes are on the left hand side. Numbers refer to base pairs. For each construction, lane a shows the probe alone without S_1 digestion; lane b, S_1 nuclease digested probe; lane c, the relevant oligonucleotide containing probe hybridized to the *in vitro* β -globin transcript and lane d the result of hybridizing the probe to the expressed RNA followed by digestion with S_1 nuclease. When the probe can form substantial secondary structure as in the case of glo² and glo⁴ (lanes c) the pattern of S_1 nuclease digestion (multiple lengths) confirms that such structures are being formed. Lane b for glo² (*) is at the far right hand side of the figure. The autoradiographs were scanned with a laser densitometer after exposure to pre-flashed films (Kodak X-AR5) with an intensifying screen. Lighter exposures than those shown were scanned, where the absorbance was proportional to the dose.



IVS-2 is mutated⁴; the *E1a* oligonucleotides ('*E1a*₉₇₃' and '*E1a*₁₁₁₁') are the two alternative 5' splice sites of the adenovirus 2 *E1a* gene (at positions 973 and 1111) that are used to give two distinct mature transcripts. Insertion of more than one oligonucleotide can be used to look at the effects of position and secondary structure.

Oligonucleotides were cloned into the *Bam*HI site of a rabbit β -globin gene in pBS^{SV}.BgIII⁷ ('pBSV') and the products verified by sequencing. The genes were introduced into HeLa cells and, after transient expression, the site of splicing was determined by S_1 mapping (Figs 1b, c). The proportion of splicing at the test site compared to the reference site was estimated by laser densitometry of autoradiographs.

The results of the S_1 mapping experiments are shown in Figs 2 and 3; quantitative estimates are given in Table 1. A number of conclusions can be drawn:

(1) The competitive assay for 5' splice site sequences allows us to rank some sequences. Splicing preferences can be shifted from the test site to the reference or vice versa depending on the sequence inserted at the test site. The results with an oligonucleotide inserted in the complementary sense (*E1a*₁₁₁₁^{comp}) confirm that insertions of 16 nucleotides into the pre-mRNA structure do not disrupt splicing patterns. This argues against a significant role for mRNA degradation induced by either the presence of extra upstream sequences when the reference splice site is used or the use of the new (test) site. It is however possible that in some cases splicing to the reference site is followed by elimination of these transcripts, due to the presence of a splice signal at the test site (for example, 'cons' in Fig. 2). Natural patterns of alternative splicing could be influenced likewise. However, this would not impair the system's discrimination between sequences that promote splicing to the test site in the presence of the competing reference site and those that do not. However, if test site sequences have different susceptibilities to degradation, this could affect the comparisons among sequences that allow splicing to the test site (points (2), (3) and (6) below). Table 1 shows that the proportion of transcripts spliced to the reference site is inversely related to the degree of homology of a test site to the consensus sequence. Thus degradation is probably insignificant or selective, in which case it would seem to

reflect a further aspect of splicing activity at the test site. In practice, the abundance of transcripts that have spliced to the reference site and carry natural splicing sequences at the test site suggests that splice site-specific degradation is not a major factor.

A weaker reference site than that of β -globin may enable better discrimination among sites that are not used in the present arrangement.

(2) The consensus sequence is used to the exclusion of the reference site; none of our other single inserts is as effective. This site, which contains at each position the nucleotide found to be most abundant in natural sites³, is so much 'stronger' than the others and small differences in sequence are so important (for example, cons and glo) that it is surprising all 5' splice sites in genes that do not exhibit alternative splicing patterns do not conform exactly to the consensus. The use of less efficient sequences would seem to imply unnecessary competition with exon sequences that by chance resemble naturally-occurring splicing signals. However, such competition might facilitate the evolution of new proteins if acceptable signals arise. The use of inefficient sequences might also alter the efficiency of gene expression. One might expect that the time taken by splicing would be minimized to reduce degradation of nascent transcripts. However, if site preference is related to the rate at which splicing occurs, then the frequent use of sub-optimal sequences suggests that there is little selective pressure for rapid splicing.

(3) The consensus sequence and *E1a*₁₁₁₁ differ only in the region that is usually within the exon, and these differences are clearly significant. It has recently become common practice for 5' splice site listings to include only the intron portion. The region included in our oligonucleotides is the region for which conserved nucleotides have been detected; it would be surprising if sequences were conserved without functional importance. We have not established whether sequences outside this region significantly influence splicing choices or efficiencies. The exon sequences may constrain the evolution of protein sequence, or the protein sequence may dictate the splice site signals; this may not be important if sub-optimal sequences are acceptable in most genes (see point 2 above), but may influence alternative splicing patterns.

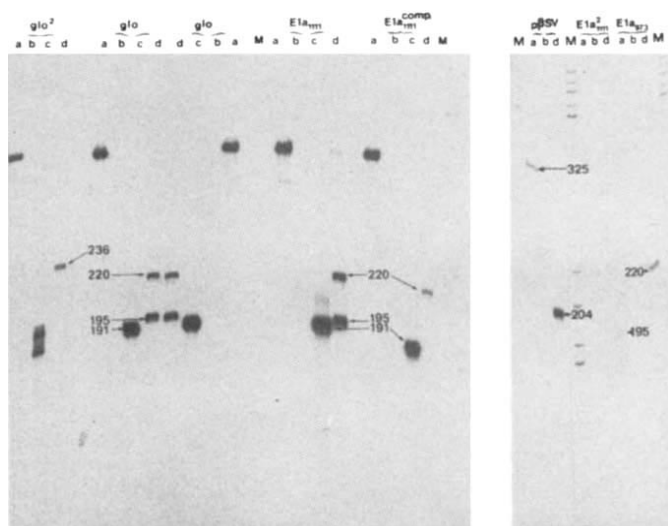


Fig. 3 Determination of the splice sites chosen using S_1 nuclease, for β -globin genes containing the following test sequences: rabbit β -globin IVS-2 5' splice site (glo), adenovirus 2 $E1a$ sites 973 and 1111 ($E1a_{973}$ and $E1a_{1111}$), a tandem dimer of $E1a_{1111}$ ($E1a_{1111}^2$) and a complementary insert ($E1a_{1111}^{comp}$), and no insert (pBSV). Lanes a, b, c, d and M are as described for Fig. 2. Numbers refer to sizes in base pairs. Samples were run on two 5% sequencing gels, containing 20% formamide; RNA was hybridized to the probe overnight at 49 °C and standard conditions were used for S_1 nuclease digestions¹⁴. The two sets of lanes for glo represent independent isolates.

(4) The sequences of the cryptic 5' splice sites are not used; thus, the fact that these sites are not used in the β -globin gene (unless the normal IVS-2 5' splice site is not available⁴) can be attributed to their sequence without invoking effects of secondary structure or position.

(5) The results from constructions glo² and glo⁴ illustrate the importance of secondary structures. Short inverted repeats can sequester a site and cause splicing to occur only at the reference. This is a new way in which secondary structure can influence splicing, involving the site directly. It has been shown¹ that an exon flanked by inverted repeats can be sequestered such that it is excised within the intron. It is possible that the effect of secondary structure would be minimized if no choice was available, and we are testing this by inactivating the reference site in glo².

(6) Another factor which affects the choice of sites is that of relative position. If we assume that the different flanking sequences of the two signals in the $E1a_{1111}^2$ tandem dimer have no bearing on their use, it seems that the upstream site is used preferentially and use of the reference site is drastically reduced. This construction gave rise to a comparatively faint band on S_1 analysis (Fig. 3); whether this is significant can be resolved only by experiments *in vitro*. The significance of linear order may be influenced by the distance between two sites or by their environment; when a single β -globin sequence is inserted it is preferred to the reference β -globin site by a ratio of only 2:1.

(7) The two $E1a$ splice site sequences are used with very different efficiencies, even though both sites are used when intact $E1a$ genes are expressed in HeLa cells. Experiments in which β -globin restriction fragments were duplicated² showed that an upstream site is usually, but not always, used to the exclusion of a downstream site. The results described in (6) emphasize the linear 5' to 3' order of alternative sites. Thus, we anticipated that in the $E1a$ gene a much more efficient sequence might be found at the downstream site in order to balance its unfavourable position. This prediction was shown to be correct (Table 1, and unpublished results with sites intermediate in efficiency between glo and cryptic sites in our assay), but further experiments are

required to test the importance of linear order and pre-mRNA structure.

(8) The fact that splicing can be directed to the inserted oligonucleotide, away from a viable and efficient natural site, has mechanistic consequences. If two 5' splice site-binding complexes^{8,9} bind simultaneously to the two sites, it is unlikely that the two sites could still be distinguished on the basis of their sequences in subsequent steps. However, we have shown here that primary sequence alone is the basis of discrimination. It is possible that both 5' sites independently bind a complex and that the ratios of site usage are related to the average lifetime of bound complex. We propose a model in which only one 5' splice site complex can bind, perhaps limited by a requirement for the presence of a 3' site complex (in agreement with some^{8,10} but not all⁹ *in vitro* data). The 5' splice site complex can bind to either 5' site, but the dissociation rates differ and the resulting pattern of usage then depends on the lifetime at each site. The involvement of only one 5' complex is rendered very likely by the proximity of our sites; experiments that increase the separation between the two 5' splice sites will reveal whether the sequence discrimination and the model only apply to such close sites.

We thank Mr J. Keyte for synthesizing the oligonucleotides, and the members of the laboratory for critically reading the manuscript. This work was supported by the MRC; J.P.E. is a recipient of an SERC postgraduate training award.

Received 27 May; accepted 24 September 1986.

1. Solnick, D. *Cell* **43**, 667-676 (1985).
2. Kühne, T., Wieringa, B., Reiser, J. & Weissman, C. *EMBO J.* **2**, 727-733 (1983).
3. Mount, S. M. *Nucleic Acids Res.* **10**, 459-472 (1982).
4. Wieringa, B., Meyer, F., Reiser, J. & Weissmann, C. *Nature* **301**, 38-43 (1983).
5. Montell, C. & Berk, A. J. *Nucleic Acids Res.* **12**, 3821-3827 (1984).
6. Montell, C., Fisher, E. F., Caruthers, M. H. & Berk, A. J. *Nature* **295**, 380-384 (1982).
7. Grosveld, G. C., de Boer, E., Shewmaker, C. K. & Flavell, R. A. *Nature* **295**, 120-126 (1982).
8. Frendewey, D. & Keller, W. *Cell* **42**, 355-367 (1985).
9. Grabowski, P. J., Seiler, S. R. & Sharp, P. A. *Cell* **42**, 345-353 (1985).
10. Perkins, K. K., Furneaux, H. M. & Hurwitz, J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 887-891 (1986).
11. Estibeiro, J. P. (in preparation).
12. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
13. Auffray, C. & Rougeon, F. *Eur. J. Biochem.* **107**, 303-314 (1980).
14. Favaloro, J., Treisman, R. & Kamen, R. *Meth. Enzym.* **65**, 718-749 (1980).
15. Sanger, F. & Coulson, A. R. *FEBS Lett.* **87**, 107-110 (1978).

Errata

Generation of single-stranded T-DNA molecules during the initial stages of T-DNA transfer from *Agrobacterium tumefaciens* to plant cells

S. E. Stachel, B. Timmerman & P. Zambryski
Nature **322**, 706-712

In this article, the title given on the contents page was incorrect. The title is correct as above.

A molecular link between the bats of New Zealand and South America

E. D. Pierson, V. M. Sarich, J. M. Lowenstein,
M. J. Daniel & W. E. Rainey
Nature **323**, 60-63

In this paper, a line was omitted from the legend for Table 1, and *yumanensis* was misspelt. The correct version reads:

The taxa used in these comparisons represent all the families with which *Mystacina* has been associated: *noctilio albiventris* (Noc) (Noctilionidae); *Mormoops megalophylla* (Mor) and *Pteronotus parnallii* (Moemoopidae); *Glossophaga soricina* (Glo) and *Phyllostomus discolor* (Phyllostomidae); *Antrozous pallidus* (Ant) and *Myotis yumanensis* (Vespertilionidae); *Tadarida brasiliensis* (Tad) (Molossidae); *Balantiopteryx plicata* (Bal) (Emballonuridae).

Immunoglobulin-derived drugs

Gary S. Hahn

Immunoglobulins may provide a novel source of drugs for the treatment and suppression of allergic reactions and other immunologically-mediated diseases.

RECENT research has suggested the possibility that small peptides derived from immunoglobulins (Igs) can serve as therapeutic drugs. All Igs secreted by plasma cells perform their effector and regulatory functions by a common pathway; by binding to Fc receptors (FcRs) in an isotype-specific manner. The precise FcR binding sites of Igs are unknown except for staphylococcal Protein A. Conventional dogma asserts that Ig FcR binding sites consist of large clusters of amino acids that require a specific three-dimensional configuration to bind to FcRs effectively. No small Ig-derived peptides with high affinities for FcRs have been reported to date. This historically led to the tacit assumption that small peptides from Igs would probably not be pharmacologically 'useful' since their affinity for FcRs would be too low.

Free-floaters

An alternative hypothesis, however, suggests that the Ig receptor binding sites consist of amino-acid clusters which also contain biologically active peptide segments. When released as free peptides, these small biologically active segments, or active site peptides (ASPs), have a low affinity for FcRs compared to intact Igs, and indeed may act as receptor agonists only after they are enzymatically released as free peptides (Fig. 1). According to this hypothesis, ASPs are released by proteolytic action while the Ig is bound to an FcR at the site of a developing immune response, and then act as short range immunoregulatory hormones.

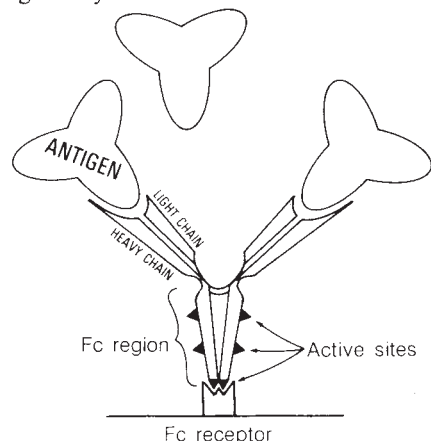


Fig. 1 Schematic of an antibody bound to two identical antigens and to an Fc receptor. The antibody's single Fc region contains multiple active site peptides (filled triangles) that bind to the Fc receptor and possibly to other immunoregulatory receptors.

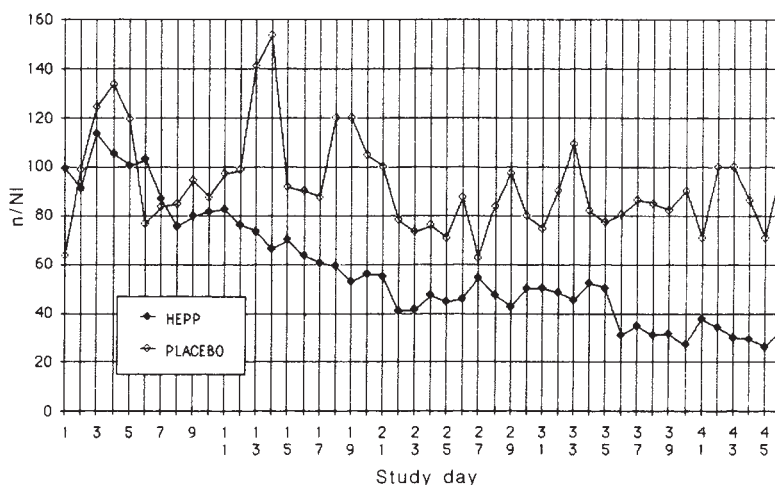


Fig. 2 Normalized stuffy nose intensity index (nIN1) during treatment with HEPP or placebo in 96 allergic rhinitis patients.

A stumbling block may exist for the concept of Ig-derived ASP agonists though, as Igs are classically thought to trigger regulatory and effector functions only when two or more Igs cross-link multiple FcRs. This mechanism seems to exclude the possibility of FcR activation by small peptides, since the proposed peptides are too small to mimic FcR cross-linking. A critical point to consider in overcoming this conceptual hurdle is that the target receptors for ASPs may not necessarily be FcRs, but may instead be different immunoregulatory receptors. A low affinity FcR-binding peptide excluded as a potential ASP by conventional dogma is therefore an ideal candidate for an ASP according to these new concepts.

Small Ig-derived ASPs could have great potential as therapeutic drugs because, unlike Igs, ASPs are likely to be non-allergenic and inexpensive to manufacture. Most importantly, ASPs would probably be rapidly degraded to amino acids and would thus lack toxicity due to long-lasting metabolites.

The first ASP derived from any Ig was the tetrapeptide tuftsin isolated from enzymatic digests of human IgG. Tuftsin was shown to stimulate phagocytosis and other protective immune responses. As monomeric IgG does not stimulate phagocytosis, it was suggested that tuftsin must first be enzymatically released before expressing activity¹.

Subsequent studies have demonstrated the existence of other Ig-derived peptides whose immunoregulatory activities appear to be unrelated to traditional Fc/

FcR interactions. The tetrapeptide rigin, and an unnamed decapeptide containing the rigin sequence, stimulate phagocytosis like tuftsin². In addition, a 23-amino acid peptide enzymatically released from human IgG, containing both the rigin and the decapeptide sequences, was reported to stimulate T cell-dependent antibody synthesis with suboptimal concentrations of antigen². A nonapeptide from the C terminus of human IgE (igercin) is also claimed to have kinin-like effects after enzymatic release, whereas an enzymatically released pentapeptide from IgM (IPM-5), located at a site in IgM spatially related to the rigin site in IgG, is claimed to exhibit thymic hormone-like activity³.

Several ASPs have demonstrated therapeutic activity *in vivo*. Tuftsin given to mice with neoplastic tumours has been shown to increase the survival rate, and the 23-amino acid human IgG peptide mentioned above was reported to augment natural killer cell activity in mice. A short peptide from human IgG (73/He.2.80) substantially prevented experimental allergic encephalomyelitis in mice, as judged by neurologic deficit scores and by the reduction of central nervous system lesions, and is expected to be tested in human clinical trials on patients with multiple sclerosis⁴.

Only one Ig-derived peptide, human IgE pentapeptide (HEPP), has been shown to have therapeutic activity in humans. HEPP's ability to inhibit IgE-mediated inflammation in humans was initially suggested to be due to competitive antagonism for the IgE FcR⁵. This

concept was later proven untenable because of HEPP's low affinity for IgE FcRs. Figure 2 shows the effect of HEPP on 'stuffy nose' during a seven-week, double-blind placebo-controlled trial in 96 allergic rhinitis patients, in which the subjects recorded daily symptoms for one week prior to treatment and for six weeks during treatment (days 8–46)⁶. As an Ig-derived ASP, HEPP's activity is unexpected based on traditional concepts of Ig/FcR interactions.

The potential identity of some of the receptors for which Ig-ASP's may act as agonists is suggested by tantalizing sequence similarities of certain ASPs and other biologically active molecules. IPM-5, for example, is identical except for one amino acid to thymopentin, an ASP of the thymic hormone thymopoietin³. The first three amino acids of IPM-5 (Arg-Pro-Asp) are nearly the same as the binding site of fibronectin and its molecular relatives vitronectin, laminin and fibrinogen (Arg-Gly-Asp)⁷. The IgE-derived igercin resembles the inflammatory peptide bradykinin³. IgG-derived tuftsin resembles the N terminus of Substance P (SP-4) and has been shown to bind to Substance P receptors, and when injected intracerebroventricularly, tuftsin induces analgesia (Substance P antagonism). SP-4 has similarly been demonstrated to have tuftsin-like phagocytosis-stimulating activity¹.

New horizons

While certainly speculative, these curious resemblances suggest the possibility that Igs may have various releasable signals coded within their structures that act as short range hormones upon non-FcR systems. This greatly challenges the conventional concept of Ig structure and function, and demands explanations for the observed biological activities of ASPs that seem unrelated to known Fc/FcR interactions. The exciting possibility that Igs may serve as design templates for novel classes of immunoregulators may serve to redirect the thinking of molecular immunologists whose initial research may have revealed only a small fraction of the potential of Ig-derived peptides to serve as therapeutic drugs. □

Gary S. Hahn is an assistant clinical professor in the Immunology and Allergy Division of the Department of Pediatrics at the University of California, San Diego School of Medicine, and is vice president for research and development at Immunetech Pharmaceuticals, 11045 Roselle Street, San Diego, California 92121, USA.

1. Najjar, V.A. & Fridkin, M. *Ann. N.Y. Acad. Sci.* **419**, 1-272 (1983).
2. Hahn, G.S. in *Protein Abnormalities* Vol. 1 (ed. Ritzman, S.E.) 193-304 (Liss, New York, 1982).
3. Chipens, G. *Surv. Immun. Res.* **4**, 220-229 (1985).
4. Armstrong, R.M., Bray, R.S. & Armstrong, D.L. presentation to the Canadian Neurologic Society International Meeting, June 1985.
5. Homburger, R.N. *Science* **189**, 389-390 (1975).
6. Dennis, S. et al. *J. Allergy Clin. Immun.* **77**, 213 (1986).
7. Leptin, M. *Nature* **321**, 728 (1986).

Immunotech ideas

More and more monoclonals are on the market, plus the equipment to produce them.

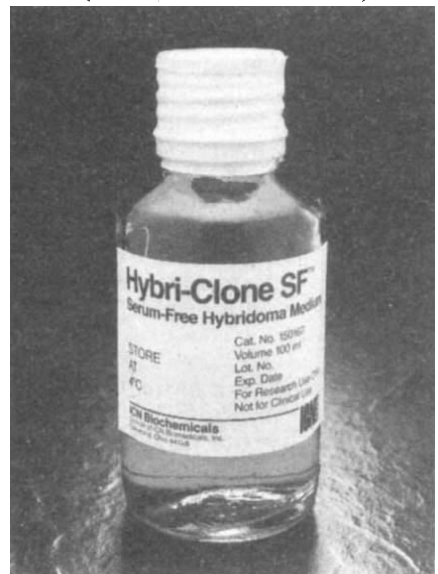
BECTON Dickinson has introduced **simultaneous two-colour immunofluorescence** with the new Simultest line of monoclonal reagents (Reader Service No. 100). Simul-



Simultest makes cells glow

test makes use of a combination of monoclonal antibodies conjugated with fluorescent dyes that are excited at the same wavelength, so that multiple tests can be run with one assay. Intended for research applications only, Becton Dickinson's reagents are currently available for identification of T cells, B cells and certain other lymphocyte populations. The cost for the Simultest T and B Cell Test is \$285 (US), and it contains enough reagent for 50 runs.

Culturing hybridomas for monoclonal antibody production requires a **serum-free medium**. ICN Biochemicals produces Hybri-Clone SF specifically for this application (Reader Service No. 101). Pro-



Serum-free for purer cultures.

vided in liquid or lyophilized form, Hybri-Clone SF is also suited for growth of primary or established cell lines. A 100 ml bottle sells for \$17.50 (US).

A Brandel **multi-probe cell harvester**, suitable for receptor binding assays, is now available from Semat Technical, Ltd. (Reader Service No. 102). Models equipped with 12, 24, and 48 wells are available, that Semat says perform cell harvesting from most sizes of test tubes within 30 seconds. Options include a hot-cold valve to separate radioactive from nonradioactive waste, a wash timer, and a collection box allowing the individual collection of filtrate. The basic 12-well M-12R model has a price tag of £2,395 (UK).

To ensure the **recovery and concentration of monoclonal antibodies** and other recombinant proteins in their biologically active forms, Sterogene has developed its ActiSept Immunoaffinity Purification System (Reader Service No. 103). The complete system includes an activated resin immunoabsorbent with a flow rate of 170 ml cm⁻² h⁻¹ (for Actigel A-Superflow) that immobilizes antigens and antibodies at physiological pH, and a nondenaturing eluent that Sterogene claims allows recovery of active proteins with a 98 per cent yield. The basic kit of 100 ml of Actigel A or Actigel A-Superflow resin plus 1 litre of elution medium costs \$270 (US). Immunogels containing anti-mouse and anti-rabbit antibodies, as well as Protein A-Actigels for the isolation of IgG, are also available.

More monoclonals

For ease in identifying cell origins during cancer and other cell culture research, Du Pont Ltd offers its Cytolite line of **monoclonal antibodies specific to cytoskeletal proteins** (Reader Service No. 104). Du Pont's monoclonals are available for vimentin, actin, tubulin and keratin, allowing experimenters to assign major cell type origins. The monoclonals typically cost £84 (UK) for sufficient antibody for 50 assays, and each order is accompanied by instructions outlining the experimental procedures for rhodamine or fluorescein detection.

Customized antibodies are available from BAbCO (Berkeley Antibody Company) (Reader Service No. 105). Its free 16-page guide outlines an extensive range of polyclonal and monoclonal antibody services. Ancillary immunological services are offered, including hapten-carrier conjugation, ELISA, and pharmacokinetic studies. BAbCO's services are keyed to researchers on any size budget, and preparations are made directly from client antigens.

MicroGeneSys, Inc. offers up-to-the-minute products for AIDS research purposes (Reader Service No. 106). Its MGStandard HIV gp160 protein is isolated from the expression of the virus envelope protein *env* in insect cells, and is thus free of contaminating mammalian cell proteins. The protein is complete except for a small deletion of approximately 100 amino acids at the C terminus, introduced for stabilization, that does not affect the two hydrophobic domains present near the C terminus of the gp41 portion. MGStandard HIV gp160 is supplied as a tested 50% pure solution in a sterile aqueous buffer at a concentration of 100 mg ml⁻¹, and a 25 mg sample costs \$385 (US). For convenience, the AIDS envelope protein precursor is also available in prepared 96-well plates for ELISA, and in impregnated nitrocellulose strips (3–4 mm × 10 mm) for Western blot analysis.

Known for radiochemicals as well as for radioligands for use in receptor detection, Amersham International has recently expanded its product line to include a range of monoclonal antibodies (Reader Service No. 107). Amersham's new receptor

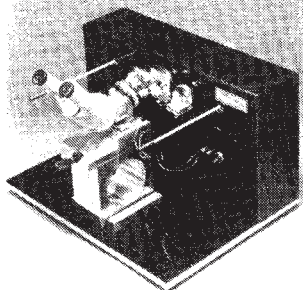


Transferrin receptors on human laryngeal carcinoma cells visualized with streptavidin-gold.

monoclonals are specific for epidermal growth factor receptor, interleukin-2 receptor, and transferrin receptor. To accompany the antibodies, Amersham also sells detection systems using biotin-streptavidin, streptavidin-gold, and DAB silver enhancement.

Flow Laboratories is now distributing Hybritech's expanded range of monoclonal antibodies, and Hybritech has recently announced three additions for research applications (Reader Service No. 108). Monoclonals for T-cell surface antigens include antibodies for interleukin-2 receptor for the detection and enumeration of activated T cells, and those for suppressor/cytotoxic and TA 5.9 T-cell antigens, useful for lymphocyte phenotyping as well as T-cell activation. Immunohistochemical detection of human, monkey and porcine chromogranin, indicating the presence of endocrine-derived cells, is

SINGER INSTRUMENTS MICRO-TOOL MAKING? then you need the SINGER MICROFORGE



Write to: Singer Instrument Co Ltd, Treborough Lodge, Roadwater, Watchet, Somerset TA23 0QL, England. Tel: (0984) 40226. Telex: 337492 Comcab G.

Reader Service No.50

Activity Monitor

Track that rat, insect, fish, limb

Tracking unit accepts input from video camera or recorder, and outputs the location of a high contrast object in the TV image. Interfaces for BBC, Apple, IBM-PC.

HVS IMAGE ANALYSING LTD,
22 Cromwell Road, Kingston,
KT2 6RE, England.



Tel: 01-546 9090

- NEW PRODUCT -
Cell-length Monitor

Reader Service No.13

Medwire Div.
stocks a wide variety of

TEFLON COATED*

Platinum alloys, stainless steel, silver, multistrand and single strand available for prompt shipment. Write for descriptive brochure and price lists.



Medwire Div.
Affiliate of Sigmund Cohn Corp
123 So. Columbus Ave.
Mt. Vernon, N.Y. 10553
914-664-5300
*Reg. T.M. of E.I. DuPont
Reader Service No.49

WESTERN BLOTS IN MINUTES

Our new GENIE electrophoretic blotter incorporates space-age plate electrodes for a uniform, high-intensity field. It is easy to load, and uses less than a liter of buffer, yet can blot gels up to 15 by 17 cm. in as little as ten minutes.

only \$190



Idea Scientific Company
P.O. Box 2078
Corvallis, Oregon 97339
(503) 758-0999

ASCB Booth No. 326

Reader Service No.53

MONOCLONAL ANTIBODIES



CELL SURFACE ANTIGENS

human T-cell T6 (CD1), clone 66-ILC7
human T-cell T3 (CD3), clone RIV9
human T-cell T3 (CD3), clone SPV-T3b
human T-cell T4 (CD4), clone RIV6
human T-cell T1 (CD5), clone 6.12.20.1 and clone 6.4.15.1
human T-cell differentiation antigen (CD7), clone WT1
human T-cell T8 (CD8), clone SPV-T8
human T-cell T8 (CD8), clone FK18
human T-cell T9/Transferrin receptor, clone 66-IIG10
human HNK-1, clone 6.13.19.1
human Monocyte/Granulocyte antigen (CD11), clone FK24
human Granulocytes (CDw15), clone FK32
HLA-D, clone 7.510.1
HLA-DOw1 related determinant, clone II B3
HLA-DRw52 related determinant, clone 7.3.19.1
human T-cell receptor, clone WT31

HORMONES

human Chorionic Gonadotropin, (hCG), clone 2092
human Follicle Stimulating Hormone, (FSH), clone 2093

INTERMEDIATE FILAMENTS

human Desmin, clone 33
human Glia Fibrillary Acidic Protein, clone 6F2
human Keratin, clone 80
human Neurofilament, clone 2F11
human Vimentin, clone 9

CYTOSKELETON ANTIGENS

actin, (rabbit muscle), clone SA1C1
tubulin, (rat), clone 1G9

PROTEINS

human Kappa, clone 2B7 and 3B10
human Lambda, clone 48
human Thyroglobulin, clone 14/14
human IgM, clone B8904 E7

HISTOLOGICAL ANTIGENS

human Endothelium, clone PAL-E
human Endothelium, clone EN4

MELANOMA ANTIGENS

Determinant discriminating between nevocellular nevi and malignant melanomas, clone PAL-M1
Determinant discriminating between nevocellular nevi and malignant melanomas, clone PAL-M2

GROWTH FACTORS

Epidermal growth factor (mouse), clone E5
Epidermal growth factor (mouse), clone F5

NEW MONOCLONAL ANTIBODIES

NK1/C3 for diagnosis of
Amelanotic Melanoma
For immuno histochemistry:
PROSTATE SPECIFIC ANTIGEN

* Polyclonal antibodies against
SEROTONINE
DOPAMINE

Other antisera against neurotransmitters
expected soon.

ENZYME LABELLED ANTIGENS.

① Toxoplasma antigen-HRP code TOX 8610
② Cytomegalovirus nuclear antigen-HRP code CMV 8620
③ Varicella-zoster virus antigen-HRP code VZV 8630

NEW PRODUCTS

SANBIO bv

Sanbio bv - biological products
Hensbergenstraat 50, 5402 EG UDEN
P.O. Box 540, 5400 AM UDEN, The Netherlands
Phone: +31 4132 - 51115 Telex: 74827 SANBO NL

Reader Service No.32

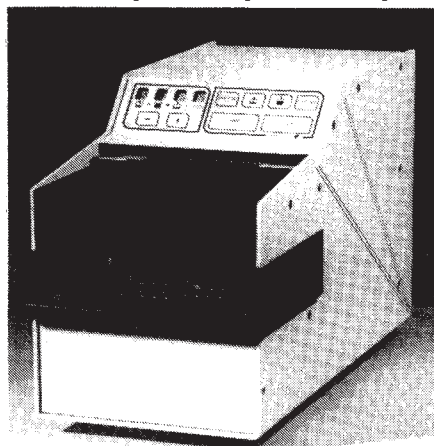
possible with Hybritech's monoclonal LK2H10. LK2H10 is also useful in Western blotting and purification of chromogranin from human adrenal glands. Monoclonal antibody to group B streptococci (Type III) is also available. Hybritech say their monoclonals are suitable for flow cytometric analysis as well as fluorescence staining, and all products are lyophilized and stable for at least one year. Prices range from £98 (UK) for a 0.2 mg sample of antibody for cell surface antigens, to £198 (UK) for 0.5 mg of antibody for tumour antigens.

Microplate machines

A microprocessor-controlled **Fluorescence Plate Reader** is available from Perkin-Elmer as an accessory to be used with its LS-3B or LS-5B fluorescence spectrometers (Reader Service No. 109). Perkin-Elmer's suggested applications for its scanner include monitoring fluorescence immunoassays and monoclonal antibody screening. The £950 (UK) plate scanner operates by using fibre optics to transmit excitation and emission energy to and from the fluorescence spectrometer, as excitation energy absorbed by the sample produces fluorescence that is detected by the emission fibres in the light pipe head.

Microplate washing and aspirating can be

tedious and time-consuming. Dynatech's Ultrawash II automates the process, and allows the operator to pre-select dispense



Dynatech's Ultrawash II is designed for those who hate to wash plates.

volumes, soak times between 0 and 90 s, rates of aspiration and number of wash cycles (Reader Service No. 110). The £2,412 (UK) machine has 96 channels, and incorporates alternative dispense/aspirate heads to allow the selection of single or multiple rows as desired.

Cambridge Life Sciences is making microtitre plate readers affordable, claim-

ing that its reader is half as expensive as the average instrument, with top quality (Reader Service No. 111). The fully portable CLS 961 is provided with a carrying case and costs £895 (UK). The CLS 961



Carry the CLS 961 anywhere.

reads 96-well ELISA and other microwell assays, and offers a choice of wavelengths of 405, 450, 492, 540, 615 and 690 nm. The more expensive CLS 962 has greater sensitivity and a wider absorbance range, and can be connected to a computer or printer.

ADVERTISEMENT

Fitting Solutions

...To your reagent needs with Becton Dickinson Monoclonal Antibodies. The largest available range of antibody formats for research and diagnostic work provides the solution, whatever the pattern of your workload.

Another example of innovative products resulting from Becton Dickinson's commitment to research and development for the clinical laboratory, the BD Monoclonal Antibody range is the widest and of the highest quality available for research and diagnostic work on leukocytes in tissue and blood. Supplied ready to use in solution, with a minimum shelf life of six months, and two years on purified immunoglobulins at 2-8 C, they combine guaranteed reliability and quality with cost-effectiveness. Convenient to use and with the best possible supporting user literature, BD MABS are used by leading scientists in the fields of AIDS, leukaemia, cancer, multiple sclerosis and transplant research.

What more fitting solution to your Monoclonal Antibody needs?

- Anti-IL2 Receptor : Activated T. cells and HTLV infected cells.
- Anti-HPCA-1 : Human progenitor cell associated antigen, reacts with acute leukaemia specimens and unipotent and multipotent progenitor cells.
- Anti - Cytokeratin (CAM 5.2) : reacts with adenocarcinomas and mesotheliomas
- Anti - Calla : human common acute lymphoblastic leukaemia antigen.

Further information available from:
Laboratory Impex Limited, Impex House,
Lion Road, Twickenham, Middlesex TW1 4JF
Tel: 01-891 4881 Telex: 8951670 Labimp G

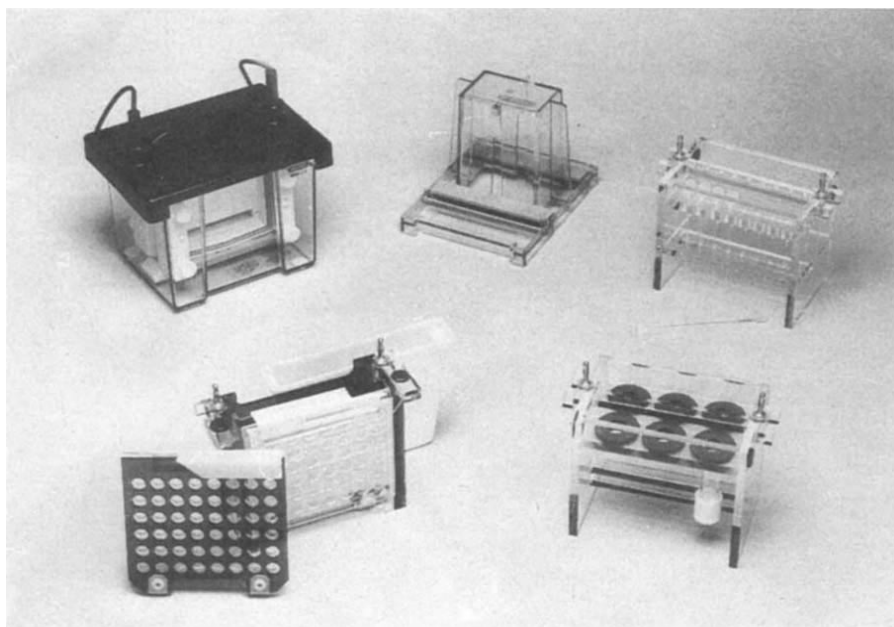
Laboratory Impex Ireland Ltd, 3A Kylesmore Park North,
Ballyfermot, Dublin 10. Tel: Dublin 262332



**BECTON
DICKINSON**

Laboratory Impex Limited

Reader Service No.6

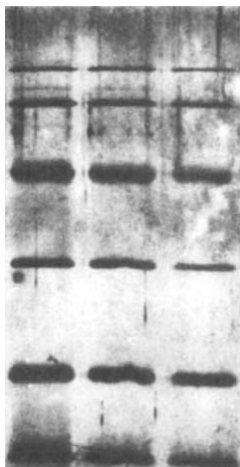


Bio-Rad's array of modular electrophoresis products.

Collecting, detecting, and dispensing

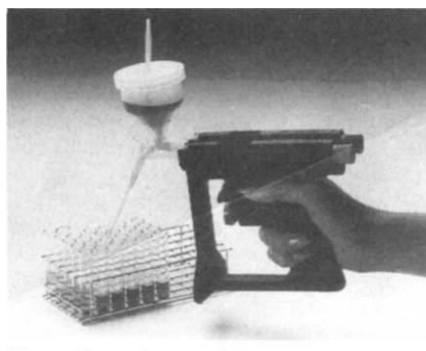
Bio-Rad has developed three **interchangeable electrophoresis modules** to expand its Mini Protean II electrophoresis cell into a complete system (Reader Service No. 112). The insertion of the Mini Trans-Blot module after the initial electrophoretic separation is completed allows electrophoretic transfers to be performed within an hour, according to Bio-Rad. By switching the module again, to the Mini Tube Cell, first dimension IEF tube gels for 2-D applications are produced. With another change of modules, the Electroeluter recovers proteins and nucleic acids from the gel. Each module may be purchased separately, or the entire system, including the Mini Protean II basic cell, is available for £1,075.50 (UK).

Bio-Rad also offers a **colloidal gold protein detection kit** for the detection of proteins bound to nitrocellulose membranes (Reader Service No. 113). The Enhanced Colloidal Gold Total Protein Detection Kit is sensitive to as little as 400 picograms of bound protein, and includes the Colloidal Gold Protein Stain, Tris, Tween-20, a Gold Enhancement Kit, and an instruction manual. The Gold Enhancement Kit deposits a layer of silver over all membrane-bound gold molecules, turning the band from red to black, and thereby increasing its visibility. The total assay time ranges from two hours to overnight. The kit has a sensitivity equivalent to silver staining of polyacrylamide gels, according to Bio-Rad, and provides protein detection blots identical in size to the immunostained membrane. Bio-Rad claims that its kit also eliminates the shrinking problem associated with anionic dyes.



Enhanced detection with gold and silver.

Manual sample dispensing and dilution is the purpose of Labsystems' Finnpiptette Diluter (Reader Service No. 114). The hand-held gun has one chamber for sample and one for diluent, and allows the sample to be aspirated, diluted, and dispensed with one fire of the trigger. The diluent chamber can be removed easily to



Shoots dispensing worries.

change reagents quickly, according to Labsystems, and is autoclavable for sterile applications. The £249 (UK) Diluter has a volume range of 0–100 µl for samples and 90–1,000 µl for reagents, and can prepare dilution ratios of between 1:1 to 1:200.

To satisfy laboratory needs associated with the increased use of radioimmunoassays (RIA), V.A. Howe & Co. Ltd. distributes a line of **Hamilton diluter/dispensers** (Reader Service No. 115). The £3,370 (UK) Microlab MT is a microprocessor-controlled four-channel dispenser designed for microtitre plate work used in ELISA. V.A. Howe also sells the smaller Microlab M for routine work, and the Microlab 1000 for high dilution ratio (1:5,000) applications.

Erratum: Zeiss microscopes

A review published in the 3 July issue of *Nature* (322, 93) incorrectly stated that "budget priced" Zeiss microscopes were available from Gallenkamp. While Gallenkamp does supply a wide range of Zeiss microscopes, the "budget" microscopes mentioned are not manufactured by Zeiss, and are badged Gallenkamp.

ADVERTISEMENTS

LIPOSOMES
REPRODUCIBLE — HOMOGENEOUS —
EXTREMELY RAPID

NEW



Our new LIPOSOMAT guarantees rapid (5ml within 60 minutes) preparation of uniformly sized unilamellar liposomes. Liposome size is selectable between 25 and ca. 600 nm diameter. Sample volume is normally 5–10ml or up to 200ml if used in combination with an additional instrument. Lipid concentration can be up to 300 mg/ml.

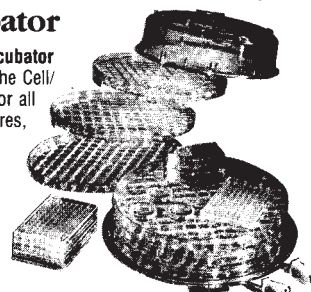
DIANORM
POB 126, D-8000 Munich 65, FRG

Reader Service No.35

Versatile Tissue Culture Incubator

Modular Incubator Chamber (the Cell/Sphere™) for all tissue cultures, aerobic or anaerobic.

Low Cost



Eliminates evaporation loss and contamination. Easy to use — flush with gas mixture, seal, place at appropriate temperature. Stackable, see-through. Quantity discounts. billups-rothenberg, inc., pob 977, Del Mar, CA USA 92014. Call 619-755-3309 or contact your Flow Laboratories representative.

Reader Service No.7

ADVERTISEMENT

Roland GRUNBERG - NANCY - FRANCE

DO
YOU
NEED
PEPTIDES?

NEOSYSTEM
LABORATORIES

PROVIDE
A COMPLETE SERVICE FOR
PEPTIDE SYNTHESIS
PEPTIDE AND PROTEIN SEQUENCING
PEPTIDE ANTIBODY PRODUCTION

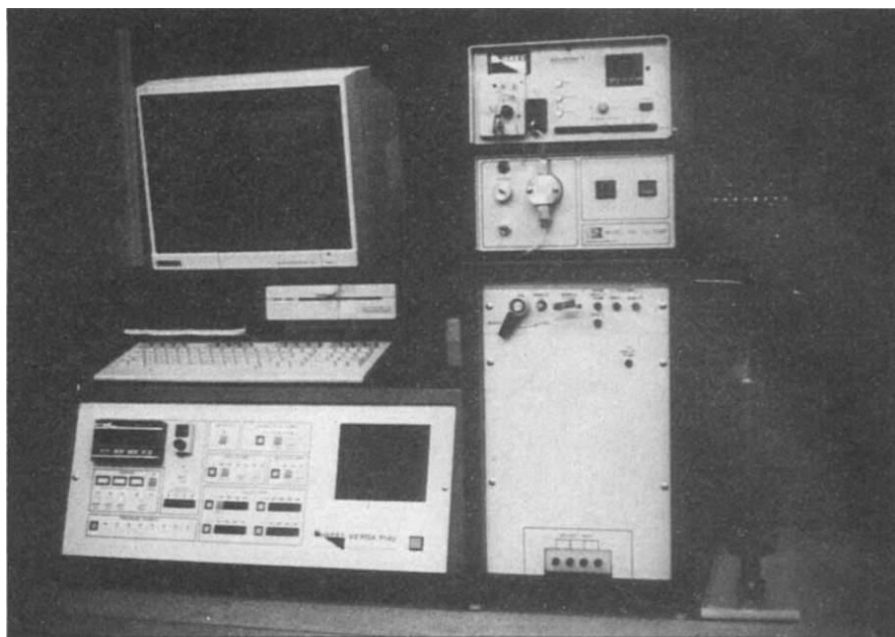
WE PROVIDE ALSO
CONTRACT RESEARCH SERVICES
TO INDUSTRIAL AND ACADEMIC
INSTITUTIONS



NEOSYSTEM
LABORATORIES

B.P. n° 309 R8
67008 STRASBOURG Cedex
FRANCE - Tél. 880 088 F
Tél. : 88.36.96.96

Reader Service No.4



The VERSAPrep-GT grows with you.

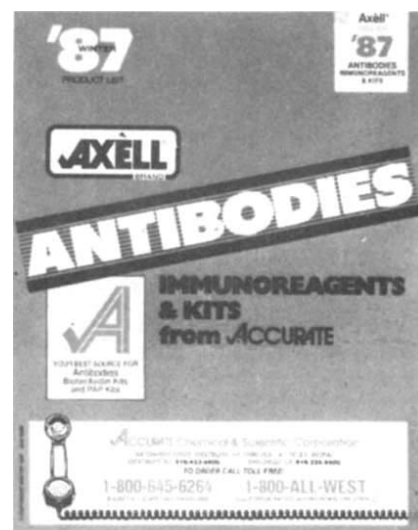
The VERSAPrep series of **modular preparative HPLC instruments** offered by Varex Corp. allows researchers to expand according to their specific needs (Reader Service No. 116). The basic VERSAPrep-GT costs \$25,465 (US) and may be operated at flow rates of up to 180 ml min⁻¹ at a maximum pressure of 3,000 p.s.i. according to Varex. Samples may be introduced with a valve-loop injector, the main eluent pump, or an accessory analytical-scale pump. Varex's HPLC offers a variety of accessory modules. A Control/Automation Module fully automates the basic VERSAPrep, while the Fraction Collection Module adds four collection ports to the VERSAPrep-GT's capabilities. The Programmable Events Module permits complete computer control, and the Column-packing Module allows the operator to slurry pack columns up to one inch in diameter. Varex says the pre-column module minimizes elution time by trapping extraneous, late-eluting components, and can be back-flushed with solvent. The introduction of the Multi-Column module expands the VERSAPrep by making it possible for two columns to be selected, either individually or sequentially, as required for a given run.

Catalogues etc.

Pharmacia is giving away a **free wall poster** describing chromatographic techniques for the purification of monoclonal antibodies (Reader Service No. 117). The 70 × 100 cm poster is designed as a useful as well as decorative addition to a laboratory, and features charge-specific, FC-region specific and immunospecific methods for chromatography and FPLC. Pharmacia's new catalogue of separation products for immunology includes infor-

mation on products for cell affinity chromatography, DNA synthesis, cell culture, cell separation by centrifugation, analytical and preparative electrophoresis systems, monoclonal antibody purification and cell studies.

Accurate Chemical has recently published a **catalogue** of its Axell Antibodies (Reader Service No. 118). The 84-page



All-purpose antibodies.

catalogue lists nearly 3000 primary and secondary antibodies, including ready-to-use diluted primary antibodies and negative controls. Accurate Chemical also provides a selection of immunochemistry kits and reagents.

These notes are compiled by Carol Ezzell from information provided by the manufacturers. To obtain further details about these products, use the reader service card bound inside the journal. Prices quoted are sometimes nominal and apply only within the country indicated.